

# MAPPING SPATIAL VARIATION OF FOLIAR NITROGEN USING HYPERSPECTRAL REMOTE SENSING

Zhihui Wang

Graduation committee:

Prof.dr. V.G. Jetten  
Prof.dr. ing W. Verhoef  
Prof.dr. K. Dabrowska-Zielinska  
Dr. J.G.P.W. Clevers

University of Twente  
University of Twente  
Institute of Geodesy and Cartography  
Wageningen University

ITC dissertation number 295  
ITC, P.O. Box 217, 7500 AA Enschede, The Netherlands

ISBN 978-90-365-4260-9  
DOI 10.3990/1.9789036542609

Cover designed by Benno Masselink  
Printed by ITC Printing Department  
Copyright © 2016 by Zhihui Wang



# MAPPING SPATIAL VARIATION OF FOLIAR NITROGEN USING HYPERSPECTRAL REMOTE SENSING

DISSERTATION

to obtain the double-badged degree of doctor  
at the University of Twente, Enschede, The Netherlands; and  
RMIT University, Melbourne, Australia,

on the authority of the rector magnificus,  
prof.dr. T.T.M. Palstra, for University of Twente

on the authority of the Chancellor, Dr Ziggy Switkowski AO; and  
the Vice-Chancellor and President of RMIT University, Mr Martin Bean CBE,  
for RMIT University

on account of the decision of the joint graduation committee,  
to be publicly defended  
on Thursday 1<sup>st</sup> of December 2016 at 11.00 hrs

by

Zhihui Wang

born on 08 March 1988

in Shandong, China

This thesis is approved by

<b>Prof. dr. A. K. Skidmore</b> , promoter	University of Twente
<b>Prof. dr. J. Hearne</b> , promoter	RMIT University
<b>Dr. T. Wang</b> , co-promoter	University of Twente

To my family



## Acknowledgements

I would like to thank a girl who decided to do a PhD four years ago, which brought me numerous wonderful adventures in the following journey. Through this long journey, I managed to find my research interests and become to what I expected. A number of elements are essential in doing a PhD, among which I value motivation, curiosity, commitment, and perservance, and more importantly the guidance and help I gained from the following people. Without their support, this thesis could not be completed.

My first gratitude goes to my promoter Prof. Dr. Andrew K. Skidmore. I was amazingly fortunate to have you as my promoter. Thank you for guiding me toward this challenging and intriguing research. I appreciate the freedom you gave me to explore on my own, as well as your enlightenment when I got stuck. You were always optimistic about my work, which boosted my confidence. From you, I learned the way of thinking critically, and that of defending my idea when it conflicted with others. Your words “work hard and work smart” echoed in my mind from time to time. I always received your critical and constructive comments and corrections on my manuscripts within short time, which dramatically improved their quality. I am indebted to you for recommending me to join the double-badged PhD program between UT and RMIT University, which greatly enriched my PhD life through the six months stay in Melbourne. Thank you for being my ideal promoter and advisor not just in academic field but also in everyday life.

Special thanks go to my daily supervisor, Dr. Tiejun Wang, who spotted my potential to do a PhD in my application nearly five years ago. The interview was the first time we discussed about doing research, which continued for four years after I started my PhD. I appreciate your guidance and support in writing the research proposal as well as later articles. I cannot remember how many times you helped me out of maize when I was writing a paper. Your logical thinking showed me how to shape my idea and tell a good story. You always replied to me promptly and gave back comments on my manuscripts within a few days. You aided me to keep on a good track. I would also like to thank you for joining my fieldwork in the first week and assisting me to set up the initial experiments. In addition, I learned from you the skills in dealing with interpersonal relationship and cooperation. Thank you for always being there to listen to me and giving me advice.

I would express my gratitude to Dr. Roshanak Darvishzadeh for your excellent supervision. Thank you for sharing your experiences of working on radiative transfer models. I also got plenty of inspirations from your publications when conducting some chapters of this thesis. Your detailed and critical comments highly improved my manuscripts. I am grateful to your efforts into advising me, also your support and kindness to me.

I am indebted to Prof. Dr. John Hearne, who was willing to be my promoter in RMIT University. I am grateful for our weekly meeting during my stay in RMIT, and appreciate our discussions on optimization and model inversion, which helped me a lot in revising one of my chapter. I would like to thank you for the corrections and comments on my manuscripts.

I am thankful to Prof. Dr. Simon Jones for providing me the working space in RMIT, sharing ideas on my research, and collaborating in one of my manuscripts. Thanks also go to Dr. Lola Suarez and Dr. William Woodgate for the discussion on my work while I was in RMIT. I appreciate the comments and experiences shared by Lola, and the help of William with processing digital hemispherical photos. I would like to thank other colleagues in the Remote Sensing Lab in RMIT who made me feel welcome, they are Phil, Mariela, Rob, Luke, and Bryan. I am grateful to the staffs Beebee, Sharon, Erin and Eliza in RMIT for their kind support.

I acknowledge the support of the “Data Pool Forestry” data-sharing initiative of the Bavarian National Park. I am indebted to the “Applied spectroscopy” team of the German Aerospace Center (DLR) and Bavarian Forest National Park for assistance in the fieldwork. Special thanks go to Dr. Uta Heiden and Dr. Macro Heurich for the contribution to my manuscripts. I appreciate that Dr. Jörg Müller patiently showed me how to operate the crossbow, was so kind when the crossbow was damaged during the fieldwork, and contacted with Dr. Martin Goßner in Technische Universität München to lend me another one, which ensured the completion of the fieldwork on time. I would also acknowledge the following people for the assistance in the fieldwork as well as collection and preprocessing of airborne hyperspectral images, they are: Dr. Nicole Pinnel (DLR), Stefanie Holzwarth (DLR), Horst Burghart (BFNP), Burkhard Beudert (BFNP), Dr. Hooman Latifi (University of Würzburg), and Dr. Martin Goßner (TUM).

I would also thank Abebe Ali for conducting the fieldwork together with me. We had worked together to overcome a variety of hardships and unexpected



circumstances both in sample collection and measurements. I appreciate your cooperation in later data analysis and all discussions on research. I would also thank Maria Buitrago for the assistance and help in the fieldwork, your efficient working impressed me a lot.

I am grateful to Esther Honderbrink for answering my inquiries, being supportive and translating the summary of the thesis into Dutch. I would like to express my gratitude to Loes Colenbrander for being kind and helpful to me, as well as formatting the thesis. Thanks go to other colleagues in NRS for their support and kindness, they are William, Thomas, Iris, Anton, Andy, and etc. I would also like to thank Benno Masselink and Job Duim for their excellent service, particularly thank Benno for the assistance in making posters and the efforts into designing the thesis cover. I express my sincere gratitude to Boudewijn de Smeth for the assistance with laboratory measurements. I would also like to thank the support of Watse Siderius in the ITC lab.

I extend my special gratitude to Linlin Li for her company and support in the past four years, we had plenty of fun in daily life as well as all types of trips. Thank all the office mates in our lady's room who are still here or has left, they are Anahita, Maitreyi, Adrie, Milica and Jing. I also thank other PhD candidates in NRS department for sharing joy and ideas, as well as the discussion in the PhD tutorial, they are Fangyuan Yu, Yiwen Sun, Jing Liu, Xi Zhu, Yifang Shi, Sonia, Elnaz, Mitra, Parinaz, Festus, Yifei Xue, Tawanda, Trini, Alby, and Haidi. I express my thanks to the Chinese friends for their support and company, they are Xiaojing Wu, Ying Zhang, Linlin Pei, Tina, Xiaoling Wang, Donghai Zheng, Xuelong Chen, Yijing Zeng, Junping Du, Xu Yuan, Hong Zhao, Xiaolong Yu, Peiqi Yang, Chengliang Liu, Shaoning Lv, and so on.

I would like to thank my family and friends in China for always being on my side. Thank Xuejun for your long-term support and company. I am most grateful to my father and mother for their tremendous support and unconditional love. I would thank my brother and sister in law who supported me and cared my parents when I was away. Lastly, I would like to thank my boyfriend Nanfeng for being my rock. Thank you for your support of each decision I made. I am sincerely grateful to you for patience, tolerance and most importantly your love.



## Table of Contents

<b>Chapter 1 Introduction .....</b>	<b>1</b>
1.1 Ecological importance of foliar nitrogen .....	2
1.2 Physical basis for spectroscopic estimates of foliar nitrogen .....	2
1.3 Hyperspectral remote sensing of foliar nitrogen.....	3
1.4 Challenges in hyperspectral remote sensing of foliar nitrogen.....	7
1.5 Objectives and research questions .....	9
1.6 Outline .....	10
<b>Chapter 2 Leaf Nitrogen Content Indirectly Estimated by Leaf Traits</b>	
<b>Derived from the PROSPECT Model.....</b>	<b>11</b>
<i>Abstract</i> .....	12
2.1 Introduction.....	13
2.2 Materials and Methods.....	15
2.3 Results.....	22
2.4 Discussion.....	27
2.5 Conclusions.....	30
<b>Chapter 3 Applicability of the PROSPECT Model for Estimating Protein</b>	
<b>and Cellulose + lignin in Fresh Leaves .....</b>	<b>33</b>
<i>Abstract</i> .....	34
3.1 Introduction.....	35
3.2 Materials and Methods.....	38
3.3 Results.....	46
3.4 Discussion.....	57
3.5 Conclusions.....	61
<b>Chapter 4 Canopy foliar nitrogen retrieved from airborne hyperspectral</b>	
<b>imagery by correcting for canopy structure effects .....</b>	<b>63</b>
<i>Abstract</i> .....	64
4.1 Introduction.....	65
4.2 Materials and Methods.....	67
4.3 Results.....	76
4.4 Discussion.....	84
4.5 Conclusions.....	87
<b>Chapter 5 Mapping leaf and canopy nitrogen content from airborne</b>	
<b>hyperspectral imagery using coupled leaf-canopy radiative transfer</b>	
<b>models.....</b>	<b>89</b>
<i>Abstract</i> .....	90

5.1 Introduction.....	91
5.2 Materials and methods.....	94
5.3 Results.....	103
5.4 Discussion.....	108
5.5 Conclusions.....	111
<b>Chapter 6 Synthesis: Hyperspectral remote sensing of foliar nitrogen...</b>	<b>113</b>
6.1 Introduction.....	114
6.2 Leaf nitrogen content indirectly estimated by leaf traits derived from the PROSPECT model.....	116
6.3 Retrieval of nitrogen by inversion of a leaf radiative transfer model PROSPECT.....	117
6.4 Nitrogen detectable from canopy BRF correcting for canopy structural effects.....	121
6.5 Mapping leaf and canopy nitrogen content by inversion of coupled leaf-canopy radiative transfer models .....	125
6.6 Convergence of leaf properties and canopy properties for foliar nitrogen retrieval .....	129
6.7 Broader implications of nitrogen estimation for ecological studies .....	132
6.8 Future research.....	133
<b>Bibliography .....</b>	<b>135</b>
<b>Summary .....</b>	<b>161</b>
<b>Samenvatting .....</b>	<b>163</b>
<b>Biography .....</b>	<b>165</b>
<b>ITC Dissertation List.....</b>	<b>166</b>

## **Chapter 1**

### **Introduction**

## **1.1 Ecological importance of foliar nitrogen**

Nitrogen is often a limiting factor for plant growth, and is important in terrestrial ecosystem carbon dynamics which acts as potential climatic feedback (Lamarque et al. 2005a; Liu et al. 2005; Magnani et al. 2007; Sievering et al. 2000). It is also an important input variable of ecosystem process models (Ollinger and Smith 2005; Plummer 2000; Smith et al. 2002). Foliar nitrogen is a primary regulator of physiological processes such as photosynthesis, leaf respiration, and transpiration (Evans 1989; Field and Mooney 1986; Reich 2012; Reich et al. 1998; Reich et al. 2006) and is related to canopy and stand-level traits such as light use efficiency, wood growth and net primary production (Ollinger and Smith 2005; Scott Green et al. 2003; Smith et al. 2002). Acknowledging the significant role of leaf nitrogen in biodiversity and ecosystem functioning, leaf nitrogen content has also been proposed as one of the essential biodiversity variables by the remote-sensing and ecology communities for satellite monitoring of progress towards the Aichi Biodiversity Targets (Pereira et al. 2013; Skidmore et al. 2015).

## **1.2 Physical basis for spectroscopic estimates of foliar nitrogen**

Nitrogen is only a relatively small constituent (0.2-6.4% dry mass) in leaves (Wright et al. 2004), which mainly exists in proteins and chlorophylls in the leaf cells. 30 - 50 % of the nitrogen in green leaves is found in a single protein, ribulose-1,5-biphosphate carboxylase-oxygenase (rubisco) (Elvidge 1990), which relates to the carboxylation of RuBP in the initial stage of CO<sub>2</sub> fixation (Collatz et al. 1992; Long 1991). Nitrogen is also present in leaf chlorophylls (6.5% by weight) (Kokaly et al. 2009), which harvest light in the photosynthesis process.

The basis of spectroscopic estimates of foliar nitrogen lies in the spectral absorption features associated with chlorophyll and protein through the visible and infrared regions of the electromagnetic spectrum. The distinct absorption properties of chlorophyll and protein lead to the development of algorithms that use both visible and near-infrared absorption features to estimate nitrogen concentration (Kokaly 2001). The strong absorption features of chlorophyll locates in the visible region 400 -700 nm due to electron energy transitions (Danks et al. 1983). The peak absorption features locate at 430 nm and 660 nm for chlorophyll a, and occur at 460 nm and 640 nm for chlorophyll b. The organic compounds absorb infrared radiation at fundamental stretching frequencies primarily in the middle infrared part of the spectrum. The absorption features

observed in the near and shortwave infrared arise from the harmonics and overtones of the fundamental stretching frequencies of C-H, O-H and N-H bonds as well as the combination bands (Hergert 1971; Osborne and Fearn 1986). Researchers with the U.S. Department of Agriculture pioneered in spectroscopic measurements of dried and ground leaves in the 1960s and 1970s mainly for assessing food and forage quality (Marten et al. 1985; Norris et al. 1976). Particular narrow band features related to nitrogen were summarized through laboratory and field studies on dried ground leaf material, which mainly locate at 910 nm, 1020 nm, 1510 nm, 1730 nm, 1980 nm, 2060 nm, 2130 nm, 2180 nm, 2240 nm and 2300 nm (Curran 1989).

The moderately strong correlation (average Pearson correlation coefficient  $r = 0.65 \pm 0.15$ ) between nitrogen and chlorophyll within and across ecosystems (Baret et al. 2007; Homolova et al. 2013; Mutanga and Skidmore 2007; Oppelt and Mauser 2004) has provided a foundation for estimating nitrogen through chlorophyll (Evans 1989; Field and Mooney 1986; Homolova et al. 2013; le Maire et al. 2008; Ustin et al. 2004). The effects of increasing nitrogen were studied on the chlorophyll absorption feature centered near  $0.68 \mu\text{m}$  (Mutanga et al. 2003). The chlorophyll absorption was deepened and widened due to increased fertilization with nitrogen before becoming saturated (Mutanga et al. 2003). A physical basis for spectroscopic estimates of nitrogen in dried and ground leaves was provided by nitrogen-containing protein absorptions (Kokaly 2001). The increase of nitrogen changes the leaf reflectance of dry leaves in the NIR absorption feature centered at  $2.1 \mu\text{m}$  (Kokaly 2001). The changes have been shown to be caused by two absorption features at  $2.055$  and  $2.172 \mu\text{m}$  that are situated on the shoulders of the  $2.1 \mu\text{m}$  absorption corresponding in wavelength position with the absorption features of proteins (Kokaly 2001).

### **1.3 Hyperspectral remote sensing of foliar nitrogen**

Despite the small quantity of nitrogen, it has been retrieved with good accuracy using leaf- and canopy-level hyperspectral data (Martin and Aber 1997; Smith et al. 2002; Wessman et al. 1988). Broad-band data often mask the spectral characteristics of biochemicals, which are typically contained in specific narrow bands, by averaging or convolving the at-sensor incoming radiation across broad wavelength ranges (Broge and Leblanc 2001; Curran 2001). Hyperspectral data is capable of detecting the narrow absorption features of nitrogen by providing contiguous, narrow spectral band information (Cho 2007a). It provides a time-

efficient and cost-effective solution to estimate foliar nitrogen compared to the traditional destructive sampling methods.

Foliar nitrogen has been determined in forest (Asner and Martin 2008a; Coops et al. 2003; Gökkaya et al. 2015a; Martin et al. 2008; Singh et al. 2015), grassland (Mutanga et al. 2004; Pellissier et al. 2015; Skidmore et al. 2010), and crop ecosystems (Inoue et al. 2012; Li et al. 2014; Yao et al. 2015). Initial research was performed using dry, ground leaf material sampled from deciduous and conifer tree species indicating the potential of spectroscopy for predicting sugar, starch, protein, cellulose, total chlorophyll, lignin and total nitrogen (Card et al. 1988). Then studies on nitrogen concentration estimation in vegetation were performed using fresh leaf spectra (Curran 1989; Kokaly and Clark 1999; Martin et al. 2008). The advent of airborne imaging spectrometer allowed the early quantification of nitrogen at canopy level using remote sensing (Peterson et al. 1988; Wessman et al. 1988). With the development of the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS), the EO-1 Hyperion space-based instrument, and numerous airborne commercial imaging spectrometers, studies have recently predicted image-scale foliar nitrogen based on more extensive plot sampling (Asner et al. 2006; Martin and Aber 1997; Smith et al. 2003; Smith et al. 2002; Townsend et al. 2003; Wessman et al. 1988).

### **1.3.1 Empirical approach**

The approaches of estimating foliar nitrogen can be basically divided into two types, viz. the empirical approach and physically based approach (Skidmore 2002). The empirical approach focuses on building statistical regression models from the high correlations between nitrogen and reflectance or its transformation (Curran 1989; Kokaly and Clark 1999; Skidmore et al. 2010; Zhao et al. 2013). Empirical techniques are dominant in nitrogen retrieval using hyperspectral data, ranging from vegetation indices (Miphokasap et al. 2012), traditional regression techniques such as stepwise multiple linear regression (Kokaly and Clark 1999), partial least square regression (Martin et al. 2008), to a number of artificial intelligence methods such as support vector regression, neural network and Bayesian model averaging (Axelsson et al. 2013; Skidmore et al. 2010; Zhao et al. 2013).

The empirical approach provides a preliminary understanding of the relation between foliar nitrogen and reflectance as well as being relatively easy to perform. However, some authors reported difficulties when transferring the empirical relationships across species or sites (LaCapra et al. 1996; Martin and Aber 1997).



Moreover, the wavebands selected by statistical analysis from different studies have, on occasion, been inconsistent and have deviated from known nitrogen absorption bands (Curran et al. 2001; Huang et al. 2004; Jacquemoud et al. 1996).

### 1.3.2 Physically based approach

The physical-based approach attempts to model the physical relationship based on radiation transfer theory. Radiative transfer models (RTMs) describe the interaction of solar radiation in leaves based on laws of optics, thus offer advantages in the robustness and transferability compared to empirical models (Darvishzadeh et al. 2011; Jacquemoud and Baret 1990; Schlerf and Atzberger 2006). RTMs have been applied to retrieve leaf biochemical parameters (e.g., leaf chlorophyll, dry matter and water content) from remotely sensed data (Baret and Fourty 1997a; Colombo et al. 2008; Darvishzadeh et al. 2008; Feret et al. 2008; le Maire et al. 2004; Omari et al. 2013; Riano et al. 2005; Zhang et al. 2008). However, there has been limited research focused on the utility of physically based models for estimating leaf nitrogen content (Dawson et al. 1998; Jacquemoud and Baret 1990; Jacquemoud et al. 1996).

Leaf radiative transfer models, *e.g.* PROSPECT, LIBERTY and LEAFMOD have been established for leaf reflectance and transmittance by modeling the radiation transfer and interaction within a leaf (Dawson et al. 1998; Ganapol et al. 1998; Jacquemoud and Baret 1990). These models simulate leaf reflectance and transmittance first, and are then used to retrieve leaf biochemical concentration via inversion. PROSPECT was proposed first for broadleaf, but has been validated and widely used for conifer needle leaves (Malenovsky et al. 2006; Moorthy et al. 2008; Zarco-Tejada et al. 2004), and requires only four input parameters (*viz.* chlorophyll, leaf water, dry matter and leaf structure).

PROSPECT has been widely applied to retrieve several leaf biochemical parameters (e.g., leaf chlorophyll, dry matter and water content) from remotely sensed data (Baret and Fourty 1997a; Darvishzadeh et al. 2008; Feret et al. 2008; Omari et al. 2013), but they have not been well developed for retrieving leaf protein and cellulose + lignin. Several studies (Fourty et al. 1996; Jacquemoud et al. 1996) attempted to incorporate leaf protein, cellulose + lignin into the absorption and scattering processes in radiative transfer models, such as the PROSPECT leaf optical properties model (Jacquemoud and Baret 1990). Leaf protein and cellulose + lignin for dry leaves could be moderately well estimated ( $R^2 = 0.49 - 0.84$ ) through PROSPECT model inversion (Fourty et al. 1996; Jacquemoud et al. 1996), but so far, these parameters have not been successfully

estimated for fresh leaves using the PROSPECT model (Botha et al. 2006; Jacquemoud et al. 1996; Kokaly et al. 2009).

When using fresh leaves, leaf reflectance and transmittance are insensitive to protein because of its small percentage of the total leaf mass (Baret and Fourty 1997b; Jacquemoud et al. 1996). Moreover, the high covariance with water and other nitrogen-containing compounds, such as chlorophyll, has led to inconsistencies in retrieving leaf protein via PROSPECT inversion (Jacquemoud and Baret 1990; Jacquemoud et al. 1996; Kokaly et al. 2009). The idea of incorporating leaf protein, and cellulose + lignin into the PROSPECT model was therefore abandoned in the 1990s (Jacquemoud and Baret 1990; Jacquemoud et al. 1996), and they were represented by a more general parameter of “dry matter content” (Fourty et al. 1996). The inversion of PROSPECT for retrieving leaf nitrogen content in fresh leaves remains a challenge and an unresolved problem. LIBERTY mainly focuses on conifer modeling (Dawson et al. 1998). The LIBERTY model also attempted to incorporate leaf nitrogen concentration as a parameter for needle leaf simulations (Dawson et al. 1998), showing that it could be done but that there was strong covariance between nitrogen and other leaf properties.

Canopy reflectance models provide a mean of understanding the covariance of leaf and canopy effects in canopy reflectance (Baret et al. 1994; Jacquemoud et al. 2000). There are generally four categories of canopy reflectance models: (1) 1D turbid medium model such as SAILH (Verhoef 1984); (2) geometrical models such as Li–Strahler GO model (Li and Strahler 1985); (3) Monte Carlo ray tracing models such as DART (Gastellu-Etchegorry et al. 1996); and (4) hybrid models such as GeoSail (Huemmrich 2001), which combine the turbid medium and the geometric models.

To select a proper canopy reflectance model, two factors need to be considered (Atzberger 2000; Pinty et al. 2004). The first factor is the realism of simulations with regard to the canopy architecture description, and the second is the invertibility of the model associated with a limited number of input variables. However, the two factors are controversial, and a compromise needs to be sought. 1-D models are suitable for homogeneous canopies such as crops and grassland, and they are easy to be inverted due to the small number of input parameters, but become problematic with regard to heterogeneous medium (Darvishzadeh et al. 2008). The GO models are most suitable for discrete canopies such as open forest, but do not account for the interactions between elements due to multiple

scattering among leaves and individual canopies (Li and Strahler 1992). Ray-tracing models can accurately model the transfer of photons and their interactions within the canopies, but their inversion is complex due to a large number of input parameters and computation complexity. The hybrid models offer a compromise between the realism of simulation of canopy and invertibility. The hybrid models benefit from a combination of 1D turbid medium and geometrical models (GO), which is more close to reality and easier to be inverted. The invertible forest reflectance model INFORM (Atzberger 2000) is an example of a hybrid model, which has been successful in retrieving vegetation parameters (Ali et al. 2016a; Schlerf and Atzberger 2006, 2012; Yuan et al. 2015).

## **1.4 Challenges in hyperspectral remote sensing of foliar nitrogen**

A number of factors confound the retrieval of leaf nitrogen from leaf and canopy spectra. At the leaf level, leaf water is one of the main obstacles for leaf nitrogen estimation using the fresh leaf spectra, because of the strong absorption of water which masks the spectral features of nitrogen in the short-wave near infrared (SWIR) spectral regions (Fourty and Baret 1998; Kokaly and Clark 1999). Other constituents in leaves such as cellulose, lignin and starch also overlap with the absorption features of nitrogen in SWIR regions, because the C-H, C-O and O-H bonds existing in proteins are also present in cellulose, lignin and starch (Curran 1989; Fourty et al. 1996). This further decreased the capability of isolating the effects of nitrogen on the variation of spectra, thus increases the difficulty in nitrogen retrieval. The overlapping absorption features of nitrogen with other leaf constituents lead to inconsistent sensitive wavelengths to nitrogen variation from empirical studies (LaCapra et al. 1996; Martin and Aber 1997), and pose a challenge in calibrating the specific absorption coefficients for each constituent in the leaf radiative transfer models (Feret et al. 2008; Wang et al. 2015a).

At the canopy level, there are a number of challenges when retrieving foliar nitrogen, including again the masking by water absorption (Fourty and Baret 1998; Kokaly and Clark 1999) as well as the confounding effects arising from canopy structure, illumination/viewing geometry and background (Asner 1998; Ross 1981; Zarco-Tejada et al. 2001). Efforts have been carried out to enhance the absorption features of nitrogen and to reduce the sensitivity of the aforementioned parameters to canopy reflectance. Spectral transformation is one of the approaches, such as using first/second derivatives and log transformation of reflectance (Coops et al. 2003; Yoder and Pettigrew-Crosby 1995). Other

approaches, such as continuum removal (Huang et al. 2004; Kokaly and Clark 1999), water removal (Ramoelo et al. 2011; Schlerf et al. 2010) and wavelet analysis (Ferwerda and Jones 2006) also improved nitrogen retrieval.

The canopy structure determines the pattern of light attenuation and the distribution of photosynthesis, respiration, transpiration and nutrient cycling in the canopy (Baldocchi et al. 2004). Hyperspectral remote sensing of foliar nitrogen may be confounded by canopy structure which causes more variation in canopy bidirectional reflectance factor (BRF) (Knyazikhin et al. 2013a). The most heavily weighted bands when calibrating foliar N are located either on the shoulders of mid-infrared water absorption features or in the regions of the NIR plateau and red edge inflection point (Smith et al. 2003). These latter features are sensitive to structural properties that influence scattering (vegetation architecture) (Ollinger 2011).

Mapping forest canopy foliar nitrogen concentration (%N) at a landscape level is practical with the aid of aircraft-based imaging spectrometry (Ollinger and Smith 2005). The current lack of higher spatial resolution satellite hyperspectral sensors impedes the mapping of nitrogen at larger scales (e.g. continental or global scale), though the launch of EnMAP mission planned for 2018 may provide an opportunity and the narrower bands on Sentinel-2 defining the red edge also offer some possibilities for nitrogen retrieval (Clevers and Gitelson 2013; Clevers and Kooistra 2012; Schlemmer et al. 2013). Ollinger et al. (2008) provided a continental-scale map of foliar nitrogen by utilizing a statistical relationship between %N and canopy bidirectional reflectance factor (BRF) for near-infrared (NIR) wavelengths (800 - 850 nm) for temperate and boreal forests. The findings were encouraging for the remote sensing community because they allow the estimation of leaf nitrogen across larger extents and potentially frequent updates using data from broadband satellite data such as MODIS (Ollinger et al. 2008).

However, Knyazikhin et al. (2013a) pointed out that the positive relationship between NIR reflectance (800 – 850 nm) and canopy foliar mass-based nitrogen concentration (%N) may also be explained by the correlation between NIR reflectance and canopy structure. They proposed the use of a structural parameter, the directional area scattering factor (DASF), for characterizing the canopy structure. Estimation of canopy foliar %N has been tested by using the ratio of BRF and DASF spectra (canopy scattering coefficient,  $W_\lambda$ ) after suppressing the impact of canopy structure (Knyazikhin et al. 2013a), however, the analyses were restricted to using information from each wavelength between 423 and 855 nm

(Knyazikhin et al. 2013a; Knyazikhin et al. 2013b). The estimation of %N using the scattering coefficient has not been investigated for longer spectral wavelengths ( $> 855$  nm), where many nitrogen absorption bands are located (Curran 1989; Fourty et al. 1996; Kokaly et al. 2009).

Meantime, Ollinger et al. (2008) argued that their hypotheses were based on the biological association between nitrogen and structural traits which affect NIR scattering and reflectance. Additionally, Townsend et al. (2013) disagreed that the %N-NIR relationship is necessarily spurious, as Wright et al. (2004) and Ollinger (2011) indicated that the canopy structure and leaf properties may co-vary across plant functional types. Species and plant functional types (*i.e.* broadleaf and coniferous forest types) account for most of the variance in canopy chemistry which has been demonstrated across tropical (Asner et al. 2015; Asner et al. 2009; Asner et al. 2012; Asner et al. 2014), temperate (McNeil et al. 2008) and boreal forests (Gökkaya et al. 2015a; Gökkaya et al. 2015b) as well as Mediterranean ecosystems (Dahlin et al. 2013). The link between species and canopy biochemistry can be explained by the concept of the ‘global leaf economics spectrum’ (Wright et al. 2004), which means that the key plant traits such as leaf mass per area, specific leaf area, leaf nitrogen, leaf phosphorous, leaf lifespan, and photosynthesis fall into a spectrum across plant species, and species converge towards the mean values of the functional traits globally (Reich et al. 1997; Wright et al. 2004). However, the covariance of these functional traits has not been fully evaluated for nitrogen estimation.

## 1.5 Objectives and research questions

The aim of the study is to retrieve foliar nitrogen at the leaf and canopy level using radiative transfer models with hyperspectral remote sensing. The research questions are as follows.

- (1) How could the relationship between nitrogen and leaf traits aid in the indirect estimation of nitrogen from a leaf radiative transfer model?
- (2) Can leaf nitrogen be retrieved from fresh leaf spectra using a leaf radiative transfer model?
- (3) Is nitrogen still detectable using canopy BRF correcting for canopy structure effects?
- (4) To what extent can foliar nitrogen be predicted at the canopy level through coupled leaf-canopy models?

## **1.6 Outline**

This thesis consists of four main chapters, each addressing one of the research questions presented in section 1.5.

Chapter 2 assessed the ability of indirect estimation of nitrogen using the link with leaf traits retrieved from PROSPECT-5 model inversion. Chapter 3 evaluated the feasibility of retrieving nitrogen through recalibrating PROSPECT-5 by incorporating the effects of leaf protein and cellulose + lignin on leaf reflectance and transmittance. Chapter 4 tested if nitrogen can still be retrieved if canopy structural effects were corrected for the canopy BRF. Chapter 5 investigated to what extent nitrogen can be retrieved by linking a recalibrated PROSPECT-5 and canopy reflectance models such as INFORM.

The thesis is concluded with Chapter 6, where the findings for each research question are presented and discussed. Chapter 6 ends with an outlook and suggestions for future research.

## Chapter 2

### **Leaf Nitrogen Content Indirectly Estimated by Leaf Traits Derived from the PROSPECT Model\***

---

\* This chapter is based on: Wang, Z., Skidmore, A.K., Darvishzadeh, R., Heiden, U., Heurich, M., and Wang, T. (2015). Leaf nitrogen content indirectly estimated by leaf traits derived from the PROSPECT model. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8, 3172-3182.

## ***Abstract***

Leaf nitrogen content has so far been quantified through empirical techniques using hyperspectral remote sensing. However, it remains a challenge to estimate the nitrogen content in fresh leaves through inversion of physically based models. Leaf nitrogen has been found to correlate to leaf traits (e.g., leaf chlorophyll, dry matter and water) well through links to the photosynthetic process, which provides potential to estimate nitrogen indirectly. We therefore set out to estimate leaf nitrogen content by using its links to leaf traits that could be retrieved from a physically based model (PROSPECT) inversion. Leaf optical (directional-hemispherical reflectance and transmittance from 350-2500 nm) and leaf biochemical (nitrogen, chlorophyll, water and dry matter) properties were measured. Correlation analysis showed that the area-based nitrogen correlations with leaf traits were higher than mass-based correlations. Hence, simple and multiple linear regression models were established for area-based nitrogen using three leaf traits (leaf chlorophyll content, leaf mass per area, and equivalent water thickness). In addition, the traits were retrieved by the inversion of PROSPECT using an iterative optimization algorithm. The established empirical models and the leaf traits retrieved from PROSPECT were used to estimate leaf nitrogen content. A simple linear regression model only using retrieved equivalent water thickness as a predictor produced the most accurate estimation of nitrogen ( $R^2=0.58$ , normalized RMSE =0.11). The combination of empirical and physically based models provides a moderately accurate estimation of leaf nitrogen content, which can be transferred to other datasets in a robust and upscalable manner.



## 2.1 Introduction

Leaf nitrogen (N) is a primary regulator of physiological processes, such as photosynthesis, leaf respiration, and transpiration (Evans 1989; Field and Mooney 1986; Reich 2012; Reich et al. 1998; Reich et al. 2006), and it is related to canopy and stand-level traits, such as light use efficiency, wood growth and net primary production (Green et al. 2003; Ollinger and Smith 2005; Smith et al. 2002). Nitrogen is also a critical factor for plant growth, and it plays an important role in terrestrial ecosystem carbon dynamics, which acts as potential climate feedback (Lamarque et al. 2005a; Morford et al. 2011; Ollinger et al. 2008; Sievering et al. 2000). Nitrogen is also an important input variable of ecosystem process models (Ollinger and Smith 2005; Plummer 2000; Zaehle et al. 2014; Zhang et al. 2013).

Although leaf nitrogen has been quantified for dry and fresh leaves, from leaf to canopy level, using imaging spectroscopy (Martin and Aber 1997; Smith et al. 2002; Wessman et al. 1988), empirical models are still the dominant method used to estimate nitrogen, such as the spectral indices model (Chen et al. 2013; Clevers and Kooistra 2012; Hansen and Schjoerring 2003; Inoue et al. 2012; Miphokasap et al. 2012; Yu et al. 2013), stepwise multiple linear regression (Curran et al. 2001; Kokaly and Clark 1999; Peterson et al. 1988), partial least squares regression (Coops et al. 2003; Martin et al. 2008; Smith et al. 2002), support vector regression (Axelsson et al. 2013), and artificial neural network (Huang et al. 2004; Mutanga and Skidmore 2004; Skidmore et al. 2010). The key challenge for empirical models relates to their transferability to different areas, since they are built on a site-specific basis. Besides, when hyperspectral data are used, the selected wavebands from different studies are inconsistent, and often deviate from nitrogen absorption bands (Curran et al. 2001; Huang et al. 2004; Jacquemoud et al. 1996).

Physically based models conceptually offer advantages compared to empirical models in robustness and transferability (Darvishzadeh et al. 2011; Jacquemoud and Baret 1990; Schlerf and Atzberger 2006), and have been widely applied to retrieve leaf biochemical parameters (e.g., leaf chlorophyll, dry matter and water content) from remotely sensed data (Baret and Fourty 1997a; Colombo et al. 2008; Darvishzadeh et al. 2008; Feret et al. 2008; le Maire et al. 2004; Omari et al. 2013; Riano et al. 2005; Zhang et al. 2008). However, limited research focused on the utility of physically based models for estimating leaf nitrogen content (Dawson et al. 1998; Jacquemoud and Baret 1990; Jacquemoud et al. 1996; Kokaly et al.

2009) . Several studies attempted to incorporate leaf nitrogen into the absorption and scattering processes in the PROSPECT leaf optical properties model (Jacquemoud and Baret 1990) and found that nitrogen content for dry leaves could be moderately well estimated through PROSPECT model inversion (Fourty et al. 1996; Jacquemoud and Baret 1990; Jacquemoud et al. 1996). But so far, leaf nitrogen content has not been successfully estimated for fresh leaves using physically based models. The reason is that for fresh leaves, leaf reflectance and transmittance are insensitive to protein because of the small percentage of nitrogen in the leaf mass (Baret and Fourty 1997b; Dawson et al. 1998; Jacquemoud et al. 1996). The strong covariance with other nitrogen-containing compounds, such as chlorophyll and water, also led to inconsistencies in retrieving nitrogen via the physically based model inversion (Dawson et al. 1998; Jacquemoud and Baret 1990; Jacquemoud et al. 1996; Kokaly et al. 2009). The idea of incorporating leaf nitrogen into the PROSPECT model was therefore abandoned in the 1990s; in other words, the inversion of PROSPECT for retrieving leaf nitrogen content in fresh leaves remains as a challenge and unresolved problem.

However, the model parameters (leaf chlorophyll content,  $CHL_{area}$ ,  $\mu g/cm^2$ ; leaf mass per area, LMA,  $g/cm^2$ ; and equivalent water thickness, EWT,  $g/cm^2$ ), which are also known as common leaf traits, can be retrieved via PROSPECT model inversion with intermediate to good accuracy (Baret and Fourty 1997a; Colombo et al. 2008; Darvishzadeh et al. 2008; Feret et al. 2008; le Maire et al. 2004; Omari et al. 2013; Riano et al. 2005; Zhang et al. 2008), and have been identified as good indicators of nitrogen (Ellsworth and Reich 1993; Evans 1989; Field and Mooney 1986). For example, the moderately strong correlation (average Pearson correlation coefficient  $r = 0.65 \pm 0.15$ ) between nitrogen and chlorophyll within and across ecosystems (Homolova et al. 2013) has provided a foundation for estimating nitrogen through chlorophyll (Evans 1989; Field and Mooney 1986; Homolova et al. 2013; le Maire et al. 2008; Ustin et al. 2004). However, using chlorophyll alone has been insufficient for estimating nitrogen, because their correlation is less strong in nitrogen-rich ecosystems (Asner and Martin 2008a). A correlation between nitrogen content and LMA has also been found across species in plant science studies, due to their close linkage with the photosynthetic process (Ellsworth and Reich 1993; Niinemets 1997; Rosati et al. 2000; Takahashi et al. 2005; Wright et al. 2004).

Leaf nitrogen content and leaf traits (i.e., chlorophyll, dry matter, and water content) can be expressed on a mass or area basis (Field and Mooney 1986;

Hikosaka 2004; Lloyd et al. 2013). The mass-based parameters are important for leaf economy, as leaf mass is considered as an investment of biomass for carbon fixation (Hikosaka 2004). The area-based parameters are important from a physiological perspective because physicochemical processes relating to photosynthesis and carbon acquisition, such as light interception, CO<sub>2</sub> diffusion, and transpiration, occur as a flux per unit leaf surface area (Hikosaka 2004; Lloyd et al. 2013). LMA, or the inverse ratio of LMA - specific leaf area - (SLA), links the mass- and area-based expressions of leaf traits (Let  $LT_{\text{mass}}$  and  $LT_{\text{area}}$  denote mass- and area-based leaf traits respectively, then  $LT_{\text{mass}} = LT_{\text{area}} / \text{LMA}$ , or  $LT_{\text{mass}} = LT_{\text{area}} * \text{SLA}$ ) (Field and Mooney 1986; Lloyd et al. 2013). The strength of correlations between leaf nitrogen content and these leaf traits varies when they are expressed on a mass or area basis (Homolova et al. 2013; Sullivan et al. 2012; Takahashi et al. 2005; Wright et al. 2004). Accordingly, the correlation between both mass-based and area-based leaf nitrogen and leaf traits needs to be explored before applying the relationship to estimating nitrogen content.

The correlations between leaf nitrogen content and leaf traits (i.e., chlorophyll, dry matter, and water content) offer potential ways to estimate nitrogen indirectly from spectral measurements. To our knowledge, there are only a few spectroscopic studies focusing on using the nitrogen links to leaf traits (i.e., LMA and EWT), although its correlation with chlorophyll has been widely considered (Homolova et al. 2013; le Maire et al. 2008). Both correlations between mass-based vs. area-based leaf nitrogen content and leaf traits need to be further understood, if they are to be extrapolated to regional or global scales. Moreover, the combination of leaf traits may improve how we can explain the variance of nitrogen. This paper aims to explore the relationship between mass-based vs. area-based leaf nitrogen content and leaf traits, and to apply it to retrieve leaf nitrogen content from fresh leaf spectra combined with physically based models.

## 2.2 Materials and Methods

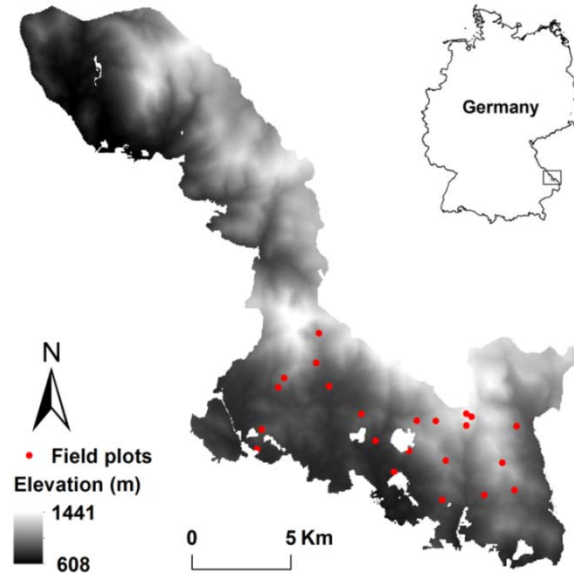
### 2.2.1 Study area and field data

The study area is located in the southern part of the Bavarian Forest National Park (49° 3' 19" N, 13° 12' 9" E), Germany (Fig. 2.1). The park has a total area of 24,218 hectares. The bedrock of the region is primarily composed of gneiss and granite. Soils weathered from these parent materials are naturally acid and low in nutrients. The main soil types are brown soils, loose brown soils and podsol

brown soils. Elevation ranges from 600 m to 1453 m. The climate is temperate with a total annual precipitation between 1200 and 1800 mm and a mean annual temperature of 5.1°C in the valleys, 5.8°C on hillsides and 3.8°C in the higher montane zones (Heurich and Neufanger 2005). Dominant species of the forests are Norway spruce (*Picea abies*) (67 %) and European beech (*Fagus sylvatica*) (24.5 %), with some white fir (*Abies alba*) (2.6 %), sycamore maples (*Acer pseudoplatanus*) (1.2 %) and mountain ash (*Sorbus aucuparia*) (3.1 %) (Heurich and Neufanger 2005). Since the mid-1990s, the forests of the National Park have been affected by massive proliferation of the spruce bark beetle (*Ips typographus*). By 2012, this had resulted in the death of mature spruce stands over an area amounting to 6000 ha (Lausch et al. 2013).

Fieldwork was conducted from mid-July to mid-August, 2013, using a stratified random sampling strategy. Land use data was obtained from Department of Conservation and Research, Bavarian Forest National Park. Based on plant functional types, the study region was stratified into broadleaf deciduous, evergreen coniferous, and mixed areas. Twenty-one plots were randomly selected from the broadleaf deciduous areas and mixed areas. Each plot was 30 m × 30 m in size, and a Leica GPS 1200 (Leica Geosystems AG, Heerbrugg, Switzerland) was used to record the center location of each plot (with an accuracy of approximately 1 m). Within each plot, depending on the homogeneity, one to three trees were selected for collecting leaf samples, resulting in overall 53 samples including 44 European beech, 4 sycamore maple, 3 mountain ash, 1 goat willow (*Salix caprea*), and 1 broad-leaved lime (*Tilia platyphyllos*). Each sample was composed of at least 20 leaves taken from the branches of an individual tree. The branches of sunlit leaves were shot down from the top of each selected tree using a crossbow. Leaves were immediately measured using a portable SPAD-502 Leaf Chlorophyll Meter (Minolta, Inc., Japan), and the averaged SPAD readings (M, unitless) for leaves in each sample were recorded.

Leaf samples were stored in zip-lock plastic bags with wet paper towels, and placed in a cooler with ice before transportation to the laboratory for further measurement.



**Figure 2.1** Location of the study area in the Bavarian Forest National Park (BFNP), Germany. The satellite image was from the world basemap in ArcGIS software (ESRI, Inc., USA).

### 2.2.2 Lab chemical analysis

The averaged SPAD value for each sample was converted to area-based leaf chlorophyll content ( $CHL_{area}$  in  $\mu\text{g}/\text{cm}^2$ ) using an empirically calibrated equation ( $\text{chlorophyll } (\mu\text{mol}/\text{m}^2) = 10^{\wedge} (M^{\wedge}0.265)$ ) provided by (Markwell et al. 1995). Though the Markwell equation was built for soybean and maize leaves, it has been tested and used for other plant species in a number of remote sensing studies (Atzberger et al. 2003; Darvishzadeh et al. 2008; Markwell et al. 1995; Yu 2014). The fresh weight ( $W_f$  in g) of leaves was measured using a digital scale, and their leaf area ( $A$  in  $\text{cm}^2$ ) was measured with a LI-3000C Portable Area Meter (LI-COR, Inc., Lincoln, Nebraska, USA). Then the spectral measurements were taken, and after that all the samples were oven-dried at  $65^\circ\text{C}$  for 48 hours, and their dry weights ( $W_d$  in g) were measured. Leaf fresh weight, dry weight and leaf area were used for deriving the water content, see details in Table 2.1. Dried samples were stored in paper bags in a cool, dark place before chemical analysis. They were then ground with a mortar and pestle to pass through a  $250 \mu\text{m}$  mesh screen. The leaf nitrogen concentration ( $N_{mass}$ , % dry weight) was determined using an AQ1 Discrete Analyzer (SEAL Analytical, Inc., Mequon, WI, USA) following a modified Kjeldahl procedure, after decomposing samples with a mixture of sulfuric acid, selenium and salicylic acid (Novozamsky et al. 1983).

**Table 2.1** The selected leaf traits in this study.

Leaf traits	Equation	Unit	References
<b>(1) Area-based leaf traits</b>			
Leaf mass per area (LMA)	$W_d/A$	g/cm <sup>2</sup>	Ellsworth et al. (1993)
Area-based leaf nitrogen content ( $N_{area}$ )	$N_{mass} \times LMA$	g/cm <sup>2</sup>	Field et al. (1986)
Equivalent water thickness (EWT)	$\rho_w \times (W_f - W_d)/A$	cm	Ustin et al. (2012)
<b>(2) Mass-based leaf traits</b>			
Mass-based leaf chlorophyll content ( $CHL_{mass}$ )	$CHL_{area}/LMA$	g/g	Field et al. (1986)
Gravimetric water content (GWC <sub>f</sub> )	$100 \times (W_f - W_d)/W_f$	%	Garnier et al. (1994)
Leaf dry matter content (LDMC)	$W_d/W_f$	mg/g	Homolova et al. (2013)

$W_f$ ,  $W_d$ ,  $A$  are leaf fresh weight (in g), dry weight (in g), and leaf area (in cm<sup>2</sup>), respectively.  $CHL_{area}$  is the area-based leaf chlorophyll content (in g/cm<sup>2</sup>).  $\rho_w$  is the physical constant representing the density of pure water (1 g cm<sup>-3</sup>). LMA and EWT are also known as area-based leaf dry matter content and leaf water content (LWC, 1 cm = 1 g/cm<sup>2</sup>), respectively, in PROSPECT (Jacquemoud et al. 1990; Ustin et al. 2012).

### 2.2.3 Spectral measurements

Leaf directional hemispherical reflectance and transmittance from 350-2500 nm were measured on 10 leaves in each sample, using an ASD FieldSpec-4 Pro FR spectrometer and an ASD RTS-3ZC Integrating Sphere designed for the spectrometer (Analytical Spectral Devices, Inc., Boulder, CO, USA). The ASD provided measured spectra with an interval of 1 nm. 200 scans per leaf were averaged to a single spectrum to minimize noise. Raw radiance was converted to reflectance using a calibrated reference standard (with approximately 99% reflectance).

The spectra were first calibrated for dark current and stray light, according to the Integrating Sphere User Manual (ASD 2008). Bands before 400 nm were removed due to noise. The spectra from ten leaves were averaged to represent each sample. They were then smoothed to minimize noise using a moving Savitzky-Golay filter (Savitzky and Golay 1964; Schmidt and Skidmore 2004) with a frame size of 13 data points (second-order polynomial). The parameters of the filter were determined to minimize noise and to maintain spectral features based on visual inspection. Only 47 of the 53 samples were considered for further analysis, because six samples became desiccated due to improper storage before the spectral measurement.

#### 2.2.4 Retrieval of leaf traits

The measurements of leaf area ( $A$ ), chlorophyll content ( $CHL_{area}$ ), nitrogen content ( $N_{mass}$ ), fresh weight ( $W_f$ ) and dry weight ( $W_d$ ) were used to derive several common area-based and mass-based leaf traits (Table 2.1).

#### 2.2.5 Statistical analysis

First, the Pearson correlation coefficient ( $r$ ) was calculated in order to investigate the association between leaf nitrogen (both mass-based and area-based) and the selected leaf traits over all samples (see Table 2.1). Species-specific correlation coefficients were not calculated due to the limited number of samples for the species other than European beech. The mass-based or area-based leaf nitrogen having in average higher correlation coefficients ( $r \geq 0.6$ ) with leaf traits were selected for later estimation. And only those leaf traits having significant correlations ( $p < 0.01$ ) with the selected leaf nitrogen were chosen for further regression analysis. Then, both simple and multiple linear regression models were developed to model leaf nitrogen content as a function of leaf traits. Unlike simple linear regression models, multiple regression models include two or three leaf traits as independent variables in order to explore if extra variables can improve the accuracy in estimating leaf nitrogen content. Non-linear regression was also tested but did not improve the estimation accuracy over linear regression models, thus linear regression models were adopted for simplicity.

Statistical tests were applied to assess the validity of the regression models, as described in (Neter et al. 1990). The model residuals (observed minus estimated value) were tested if they are independent, homoscedastic and normally distributed. F-test was performed to test the significance of the linear regression model, and t-test was performed to test the significance of individual regression coefficients. Besides, the variance inflation factor (VIF) was assessed for correlated predictor variables in the multiple regression models. Independent variables with VIF under 10 were considered as indication of low multicollinearity (O'Brien 2007).

All regression models were validated using a “leave-one-out” cross-validation, with the final model developed using all 47 samples. For each dependent variable, a model is built using independent variables from 46 samples, which is used to estimate the value of the left-out sample. The procedure is repeated for 47 times in order to obtain estimates for all samples. The cross-validated coefficient of determination ( $R^2_{cv}$ ), root mean square error ( $RMSE_{cv}$ ), and normalized RMSE

(NRMSE<sub>CV</sub>) between predicted and measured values were calculated to evaluate the performance of the regression models (Schlerf et al. 2005). The cross-validated RMSE is a good indicator of the accuracy of the model in predicting unknown samples, because the predicted samples are different from the samples used to build the model (Schlerf et al. 2005). For multiple regression models, standardized regression coefficients (beta coefficients) were calculated to compare the relative contribution of each independent variable to estimate leaf nitrogen content (Bring 1994; Landis 2005). Statistical analysis was conducted using IBM SPSS Statistics 20 (IBM, Inc.) and MATLAB (The MathWorks, Inc.).

### **2.2.6 Estimation of CHL<sub>area</sub>, LMA and EWT using the PROSPECT model**

The PROSPECT leaf optical properties model was developed to simulate leaf directional-hemispherical reflectance and transmittance over the optical domain from 400-2500 nm (Jacquemoud and Baret 1990). It only needs four input parameters: leaf structure index (N<sub>struc</sub>), leaf chlorophyll content (CHL<sub>area</sub>, µg/cm<sup>2</sup>), leaf dry matter content (known as leaf mass per area, LMA, g/cm<sup>2</sup>), and leaf water content (known as equivalent water thickness, EWT, cm, in Table 2.1). The improved (1-nm resolution) and recalibrated version, PROSPECT-4, was chosen in this study (Feret et al. 2008).

Using an iterative optimization inversion algorithm, the four input parameters of PROSPECT were estimated using measured reflectance and transmittance spectra. The inversion was performed using a bounded optimization package, FMINSEARCHBND.M in MATLAB. (The Matlab code is available in D'Errico (2012)). The inversion process was to find the parameter vector  $\theta = [N_{struc}, CHL_{area}, LMA, EWT]^T$ , which minimizes the merit function:

$$J(\theta) = \sum_{\lambda_1}^{\lambda_2} (R_{mes}(\lambda) - R_{mod}(\lambda, \theta))^2 + (T_{mes}(\lambda) - T_{mod}(\lambda, \theta))^2 \quad (1)$$

where  $\lambda$  is the wavelength,  $R_{mes}$  and  $T_{mes}$  are respectively the measured reflectance and transmittance, and  $R_{mod}$  and  $T_{mod}$  are the modeled values.

The range of the input parameters, CHL<sub>area</sub>, LMA and EWT (Table 2.3), specified in the bounded optimization function were determined based on the prior field information (Table 2.2). The range for the leaf structural index N<sub>struc</sub> was chosen based on an earlier study by (Jacquemoud and Baret 1990), which reported that dicotyledons have N<sub>struc</sub> values between 1.5 and 2.5. The model inversions included three steps, and in each step, an optimization was performed for



estimating different parameters over a different wavelength range. For each range, the parameters which were estimated have the greatest influence on the reflectance and transmittance. Details were as follows: (1) leaf structure index was determined over 760-1300 nm following the three wavelengths method described in Jacquemoud et al. (1996) and Feret et al. (2008); the accuracy of estimation for the other three parameters can be improved if the leaf structure index is computed first and used as a known value (Li and Wang 2011). (2) with the leaf structure index fixed to the value retrieved in the first step, along with prior information for ranges of LMA and EWT, the leaf chlorophyll content was estimated using 450-690 nm by inversion of PROSPECT (Asner and Vitousek 2005); (3) with the estimated leaf structure index and estimated leaf chlorophyll content, LMA and EWT were concurrently estimated over the 900-2500 nm range, where water and dry matter are most absorbent (Baret and Fourty 1997a; Feret et al. 2008; Fourty et al. 1996).

The accuracies of the retrieved parameters ( $CHL_{area}$ , LMA and EWT) via model inversion were evaluated by the coefficient of determination ( $R^2$ ), the root mean square error (RMSE), and the normalized RMSE ( $RRMSE = RMSE/range$ ) between the estimated and measured values.

**Table 2.3** Specific ranges for parameters in the PROSPECT model

Parameter	Abbreviation	Unit	Minimum value	Maximum value
Leaf structure index	$N_{struc}$	-	1.5	2.5
Leaf chlorophyll content <sup>a</sup>	$CHL_{area}$	$\mu g/cm^2$	20	60
Leaf mass per area <sup>a</sup>	LMA	$g/cm^2$	0.002	0.015
Equivalent water thickness <sup>a</sup>	EWT	cm	0.002	0.017

<sup>a</sup> The ranges were determined based on the prior knowledge from field measurements.

### 2.2.7 Estimation of leaf nitrogen content

The regression models passing all statistical tests were coupled with their corresponding retrieved leaf traits from the PROSPECT model inversion (Section 2.2.6) to estimate mass-based or area-based leaf nitrogen content (see Section 2.2.5). The coefficient of determination ( $R^2$ ), the root mean square error (RMSE), and NRMSE between the measured leaf nitrogen content and predicted values were used to evaluate the performance of each model.

## 2.3 Results

### 2.3.1 Characteristics of leaf properties

Table 2.2 summarizes the statistical characteristics of leaf properties in our dataset. It shows the area-based leaf chlorophyll content ranged from 19.47-58.91  $\mu\text{g}/\text{cm}^2$ , while the mass-based leaf chlorophyll content varied from 2.00-14.09  $\mu\text{g}/\text{g}$ . The area-based leaf nitrogen content spanned an 8-fold range of values, while the variation of mass-based leaf nitrogen content was limited (1.63-3.97%). LMA had a 5-fold range of values. EWT ranged from 0.0043-0.0156 cm. While area-based plant traits had a wide range of values, the mass-based plant traits,  $\text{GWC}_f$  and LDMC showed limited variation.

**Table 2.2** The statistics of leaf traits (sample size = 47).

	$N_{\text{area}}$ ( $\text{g}/\text{cm}^2$ )	$N_{\text{mass}}$ (% $W_d$ )	$\text{CHL}_{\text{area}}$ ( $\mu\text{g}/\text{cm}^2$ )	$\text{CHL}_{\text{mass}}$ ( $\text{mg}/\text{g}$ )	LMA ( $\text{g}/\text{cm}^2$ )	LDMC ( $\text{mg}/\text{g}$ )	EWT (cm)	$\text{GWC}_f$ (%)
Min	6.73E-05	1.63	19.47	2.00	0.0022	337.3	0.0043	45.87
Max	4.45E-04	3.97	58.91	14.09	0.0122	541.3	0.0156	66.27
Mean	2.13E-04	2.88	34.75	4.96	0.0075	460.2	0.0087	53.98
Std.	5.70E-05	0.51	8.16	1.96	0.0018	437	0.0019	4.37

See Table 2.1 for the details of leaf traits.

### 2.3.2 Correlations of leaf traits and linear regression models

Table 2.4 illustrates the relationships between leaf nitrogen and selected leaf traits for pooled samples (details in Table 2.1). Species-specific correlation coefficients calculated for the European beech were similar but slightly lower (results not shown). To involve the variations across species, the results of the pooled samples were utilized. For  $N_{\text{mass}}$ , the highest positive relationship was with  $\text{GWC}_f$  ( $r = 0.663$ ,  $p < 0.01$ ), followed by area-based leaf chlorophyll content ( $r = 0.637$ ,  $p < 0.01$ ) and mass-based chlorophyll content ( $r = 0.555$ ,  $p < 0.01$ ). A weaker correlation was observed between  $N_{\text{mass}}$  and LMA ( $r = -0.346$ ,  $p < 0.05$ ). In terms of  $N_{\text{area}}$ , we found the highest correlation with EWT ( $r = 0.841$ ,  $p < 0.01$ ), followed by LMA ( $r = 0.686$ ,  $p < 0.01$ ). Moderate correlation was observed between  $N_{\text{area}}$  and area-based leaf chlorophyll content ( $r = 0.597$ ,  $p < 0.01$ ), and no significant correlation was found between  $N_{\text{area}}$  and mass-based leaf chlorophyll content. Since  $N_{\text{area}}$  had a higher correlation with leaf traits than  $N_{\text{mass}}$ , we only considered area-based nitrogen and its highly correlated leaf traits for later analysis. Furthermore, the parameters that are incorporated in PROSPECT are also area-based, which enables the use of well-fitting relationships on the leaf traits derived from the model.

All regression models passed the statistical tests except three multiple regression models in the t-test (Table 2.5). The regression coefficient of LMA in Equation No. 4, that of  $CHL_{area}$  in No.6, and that of EWT in No. 7 (Table 2.5) were not significant. Multicollinearity among independent variables for multiple regression models ( $VIF < 10$ ) were not observed (Table 2.5).

The equations of simple and multiple linear regression models are listed in Table 2.6. The performance of each model in predicting the leaf nitrogen content ( $N_{area}$ ) is illustrated in Table 2.7. Among the simple linear regression models, the highest  $R^2$  values ( $R^2_{CV} = 0.660$ ) and lowest RMSE ( $RMSE_{CV} = 3.29E-05 \text{ g/cm}^2$ ) occurred when using EWT as the independent variable. In comparison, low  $R^2$  were obtained when regressing with LMA ( $R^2_{CV} = 0.376$ ) or  $CHL_{area}$  ( $R^2_{CV} = 0.243$ ).

The multiple linear regression model using LMA and  $CHL_{area}$  as predicting variables greatly improved the accuracy of estimation ( $R^2_{CV} = 0.695$ ) compared to simple regression models using only one of the two variables. The two variables had almost equal effects on estimating  $N_{area}$ , as seen from the standardized regression coefficients (Table 2.7, coefficients = 0.631 and 0.532 for LMA and  $CHL_{area}$ , respectively). No collinearity was found between LMA and  $CHL_{area}$ .

Other three multiple linear regression models (Equation No. 4, 6 and 7 in Table 2.6) with extra variables also provided accurate prediction of  $N_{area}$  ( $R^2_{CV} = 0.648$  - 0.710, Table 2.7). In terms of the model using EWT and LMA as predictors, EWT played the dominant role in explaining the variance of  $N_{area}$  (Table 2.7, standardized regression coefficients = 0.710 and 0.186 for EWT and LMA, respectively). This can be attributed to the correlation between EWT and LMA ( $r = 0.704$ ,  $p < 0.01$ ), which is also the reason that the regression coefficient of LMA in the model was not significant (Table 2.5). The same phenomena was observed in the model developed using EWT and  $CHL_{area}$ . The multiple regression model incorporating all three leaf traits – LMA, EWT and  $CHL_{area}$  – gave the highest  $R^2$  values ( $R^2_{CV} = 0.710$ ) and lowest RMSE ( $RMSE_{CV} = 3.04E-05 \text{ g/cm}^2$ ). All three independent variables had a similar effect on predicting  $N_{area}$ , according to their respective coefficients of 0.422, 0.324 and 0.350. These three multiple regression models (Equation No. 4, 6 and 7 in Table 2.6) were not involved for nitrogen estimation in later analysis, due to their non-significant coefficients.

**Table 2.4** Correlations between leaf nitrogen content and other leaf traits

	CHL <sub>area</sub>	CHL <sub>mass</sub>	LMA	LDMC	EWT	GWC <sub>f</sub>
N <sub>area</sub>	0.597**	-0.282	0.686**	0.066	0.841**	-0.066
N <sub>mass</sub>	0.637**	0.555**	0.346*	0.663**	0.200	0.633**

\*\* correlations significant at  $p < 0.01$ , and \* correlations significant at  $p < 0.05$ . See Table I for the definitions of leaf traits.

**Table 2.5** Statistical tests for validity of regression models for estimating leaf nitrogen content (N<sub>area</sub>) using different combinations of independent variables

Equation No.	Independent variables	$p$ -value (F-test)	$p$ -value (t-test)			VIF		
			CHL <sub>area</sub>	LMA	EWT	CHL <sub>area</sub>	LMA	EWT
1	CHL <sub>area</sub>	0.00	0.00			1		
2	LMA	0.00		0.00			1	
3	EWT	0.00			0.00			1
4	LMA, EWT	0.00		0.10	0.00		1.98	1.98
5	LMA, CHL <sub>area</sub>	0.00	0.00	0.00				
6	EWT, CHL <sub>area</sub>	0.00	0.28		0.00	1.65		1.65
7	EWT, CHL <sub>area</sub> , LMA	0.00	0.00	0.00	0.06	2.66	3.19	5.26

**Table 2.6** Regression equations between leaf nitrogen content (N<sub>area</sub>) and different combination of independent variables

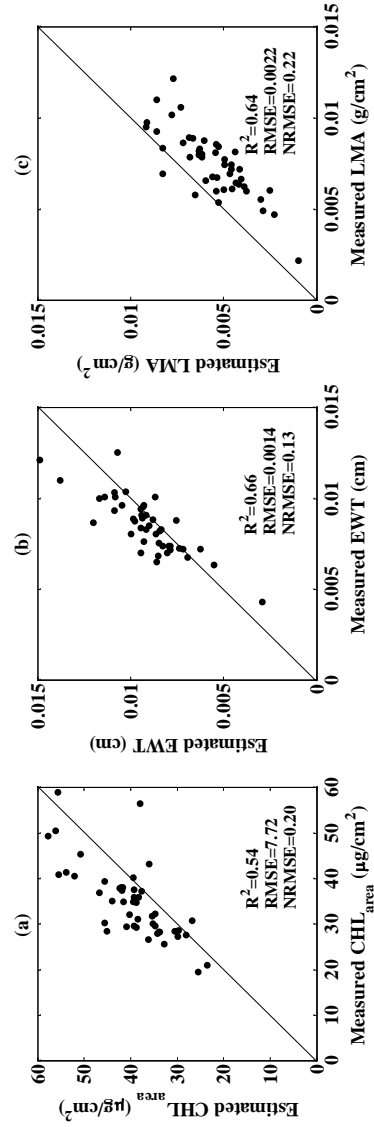
Equation No.	Regression Equations
1	$N_{area} = 4.173E-06CHL_{area} + 6.796E-05$
2	$N_{area} = 2.184E-02LMA + 4.907E-05$
3	$N_{area} = 2.563E-02EWT - 1.056E-05$
4	$N_{area} = 5.922E-03LMA + 2.164E-02EWT + 2.018E-05$
5	$N_{area} = 2.010E-02LMA + 3.718E-06CHL_{area} + 6.705E-05$
6	$N_{area} = 2.346E-02EWT + 7.922E-07CHL_{area} + 1.916E-05$
7	$N_{area} = 1.343E-02LMA + 9.873E-03EWT + 2.447E-07CHL_{area} + 5.893E-05$

**Table 2.7** Performance of regression models for estimating leaf nitrogen content ( $N_{\text{area}}$ ) using different combinations of independent variables

Equation No.	Independent variables	$R^2_{CV}$	$RMSE_{CV}$	$NRMSE_{CV}$	Standardized coefficients		
					$CHL_{\text{area}}$	LMA	EWT
1	$CHL_{\text{area}}$	0.243	4.91E-05	0.130	0.597		
2	LMA	0.376	4.46E-05	0.118		0.686	
3	EWT	0.660	3.29E-05	0.087			0.841
4	LMA, EWT	0.663	3.28E-05	0.087		0.186	0.710
5	LMA, $CHL_{\text{area}}$	0.695	3.12E-05	0.083	0.532	0.631	
6	EWT, $CHL_{\text{area}}$	0.648	3.35E-05	0.089	0.113		0.770
7	LMA, EWT, $CHL_{\text{area}}$	0.710	3.04E-05	0.080	0.350	0.422	0.324

### 2.3.3 PROSPECT model inversion for $CHL_{\text{area}}$ , EWT and LMA

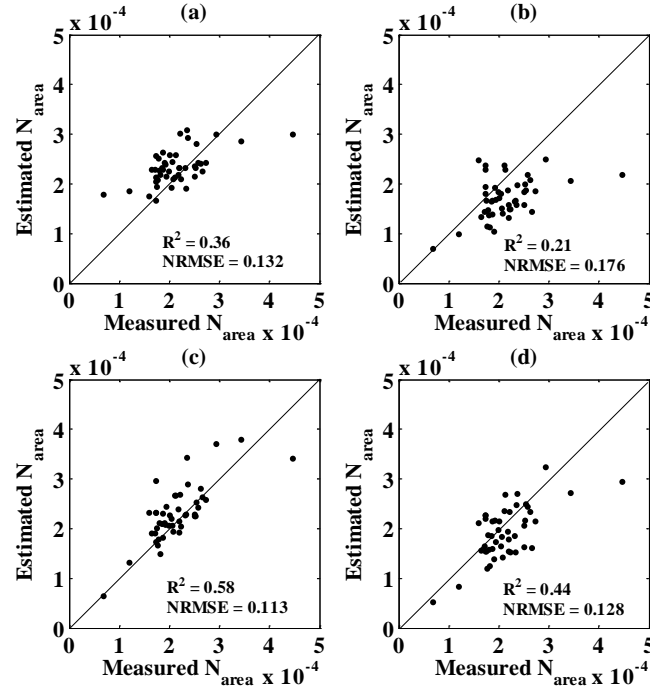
Leaf traits serving as independent variables in regression models were predicted from the PROSPECT model inversion. The inversion provided accurate estimates of  $CHL_{\text{area}}$  ( $R^2 = 0.54$ ,  $RMSE = 7.72 \mu\text{g}/\text{cm}^2$ ), EWT ( $R^2 = 0.66$ ,  $RMSE = 0.0014 \text{ cm}$ ) and LMA ( $R^2 = 0.64$ ,  $RMSE = 0.0022 \text{ g}/\text{cm}^2$ ) (Fig. 2.2). The inversion accuracy obtained here is comparable to previous studies (Baret and Fourty 1997a; Feret et al. 2008; Riano et al. 2005; Romero et al. 2012). Based on the values of normalized RMSE (Fig. 2.2), it can be noted that EWT ( $NRMSE = 0.13$ ) was estimated with a higher accuracy than LMA ( $NRMSE = 0.22$ ) or  $CHL_{\text{area}}$  ( $NRMSE = 0.20$ ).



**Figure 2.2** Measured versus estimated leaf traits obtained from PROSPECT model inversion: (a)  $CHL_{area}$ , leaf chlorophyll content, (b) EWT, equivalent water thickness and (c) LMA, leaf mass per area.

**Table 2.8** Validation of estimated  $N_{area}$  from the combination of regression models and predicted  $CHL_{area}$ , EWT, LMA from the PROSPECT inversion

Independent variables	$R^2$	RMSE ( $g/cm^2$ )	NRMSE
$CHL_{area}$	0.36	4.98E-05	0.132
LMA	0.21	6.64E-05	0.176
EWT	0.58	4.26E-05	0.113
LMA, $CHL_{area}$	0.44	4.85E-05	0.128



**Figure 2.3** Comparison between measured and estimated  $N_{\text{area}}$  ( $\text{g}/\text{cm}^2$ ) using different regression models, coupled with their independent variables retrieved from the PROSPECT model inversion. The independent variables of each regression model were (a)  $\text{CHL}_{\text{area}}$ , (b) LMA, (c) EWT, (d) LMA and  $\text{CHL}_{\text{area}}$ .

### 2.3.4 Accuracy of estimated leaf nitrogen content

$N_{\text{area}}$  was estimated with a combination of empirical regression models and the retrieved leaf traits through the physically based model inversion using PROSPECT. Three simple linear regression models (Equation No. 1-3 in Table 2.6) and one multiple regression model (Equation No. 5 in Table 2.6) were used in the estimation. Of the four models, the most accurate estimation was achieved by the linear regression model using retrieved EWT yielding an  $R^2$  of 0.58, and an RMSE of  $4.26\text{E}05 \text{ g}/\text{cm}^2$  (Fig 2.3.(c) and Table 2.8). The remaining three regression models offered less accurate estimations of leaf nitrogen, with the linear regression model using LMA providing the least accurate estimation ( $R^2 = 0.21$ , RMSE =  $6.64\text{E}-05 \text{ g}/\text{cm}^2$ , Table 2.8).

## 2.4 Discussion

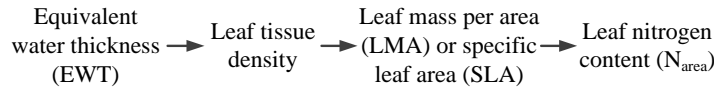
This study confirms the feasibility of estimating leaf nitrogen content by combining empirical and physically based models; it has previously only been

demonstrated using empirical methods (Kokaly et al. 2009; Martin and Aber 1997; Martin et al. 2008; Wessman et al. 1988), and had been rejected as a suitable approach for using physical model inversion with fresh leaves (Baret and Fourty 1997a; Fourty et al. 1996; Jacquemoud et al. 1996). Leaf nitrogen content ( $N_{\text{area}}$ ) was moderately well estimated indirectly through the PROSPECT model inversion using correlated leaf traits as the main driver.

Higher correlations were found between area-based leaf nitrogen content and leaf traits ( $\text{CHL}_{\text{area}}$ , LMA and EWT) than when nitrogen and leaf traits were expressed on a mass basis. This phenomena can be explained by the role of nitrogen and leaf traits in photosynthesis: most processes such as light interception and carbon acquisition are expressed on a leaf surface area basis (Hikosaka 2004; Lloyd et al. 2013). The indirect estimation of leaf nitrogen content through the PROSPECT model inversion relies on  $\text{CHL}_{\text{area}}$ , EWT and LMA, which are expressed on an area basis.

The role of leaf traits (LMA and EWT) in estimating leaf nitrogen content as a medium has been largely ignored in the literature, although chlorophyll is often considered as a proxy of nitrogen (Homolova et al. 2013; le Maire et al. 2008). In this study, a higher correlation was found between leaf nitrogen content and LMA/EWT than chlorophyll. The top-of-canopy sunlit leaves, exposed to illumination, are above the saturation level for photosynthesis (Poorter et al. 1995) and the fraction of leaf nitrogen allocated to chlorophyll becomes constant (Evans and Poorter 2001), while more nitrogen is invested in additional carbon-fixing compounds (Harrison et al. 2009; Niinemets 2010). In addition, only 19% of leaf nitrogen in C3 plants is allocated to light harvesting complexes (Evans 1983), and only 1.7% is directly in chlorophyll, whereas around 70% of nitrogen is in molecules that are related to carbon fixation (Chapin et al. 1987). LMA is a direct measure of leaf dry-mass investment per unit of light-intercepting leaf area deployed, and it gathers carbon-related compounds, such as cellulose, lignin, hemicellulose and protein (Fourty et al. 1996; Jacquemoud et al. 1996). The strong relationship between  $N_{\text{area}}$  and LMA in our study confirms these facts and is also consistent with earlier findings (Ellsworth and Reich 1993; Niinemets 1997; Rosati et al. 2000; Sullivan et al. 2012; Takahashi et al. 2005; Wright et al. 2004).





**Figure 2.4** Indirect relationship between equivalent water thickness (EWT) and leaf nitrogen content ( $N_{\text{area}}$ ).

However, it is difficult to explain the correlation between  $N_{\text{area}}$  and EWT from a physiological perspective, because leaf water content is highly dynamic in the temporal dimension compared with variation in  $N_{\text{area}}$  and LMA (Ustin et al. 2012). Firstly, from statistical analysis, the high correlation between LMA and EWT ( $r = 0.704$ ,  $p < 0.01$ ) supports the notion of a relationship between  $N_{\text{area}}$  and EWT. A similar relationship between  $N_{\text{area}}$  and EWT ( $r = 0.814$ ,  $p < 0.01$ ) was also observed in the study of six tree species, including both deciduous and coniferous species by (Sullivan et al. 2012). Secondly, EWT is related to  $N_{\text{area}}$  by way of LMA or SLA (Fig. 2.4). EWT (leaf water) is a major determinant of SLA, since increased EWT results in decreasing leaf tissue density (Garnier and Laurent 1994; Meziane and Shipley 1999). SLA is determined by leaf tissue density and leaf thickness (Garnier and Laurent 1994; Witkowski and Lamont 1991). While SLA or LMA is linked to  $N_{\text{area}}$  through the photosynthetic process, a connection between EWT and  $N_{\text{area}}$  is also expected. The last reason is possibly due to the role that leaf water plays in leaves. It provides interactions with nitrogen content in an indirect way, by transporting nutrients and acting as a regulator of photosynthesis (Asbjornsen et al. 2011; Kramer and Boyer 1995; Waring and Landsberg 2011).

Among the four regression models involved for nitrogen estimation,  $N_{\text{area}}$  was most accurately estimated with a simple linear regression model using retrieved EWT than retrieved LMA or  $\text{CHL}_{\text{area}}$ . A possible explanation is that EWT generates a model with a high explained variance of  $N_{\text{area}}$  compared with other leaf traits. Another reason is that EWT may have been retrieved with a higher accuracy than LMA, as confirmed in other studies (Clevers et al. 2008; Colombo et al. 2008). Though the multiple regression model using measured LMA and  $\text{CHL}_{\text{area}}$  gave accurate estimation of  $N_{\text{area}}$ , its performance was deteriorated when it was applied to the leaf traits retrieved from the PROSPECT inversion. This is because the error in model inversion was propagated into the regression model, and affected the accuracy of nitrogen estimation. Similar phenomena was observed in the simple linear regression models using LMA or  $\text{CHL}_{\text{area}}$  as an independent variable. Although error-propagation is true for all regression models, it is more severe for models using LMA or  $\text{CHL}_{\text{area}}$  than that with EWT, due to different accuracies of the retrieved parameters from model inversion. A

higher accuracy of predicted leaf nitrogen content is expected if LMA and  $CHL_{area}$  are retrieved more accurately from model inversion.

In the establishment of multiple linear regression models, e.g. Equation No. 4, 6 and 7 in Table 2.6, the improvement in the accuracy of predicting nitrogen was limited when adding extra variables to supplement EWT (Table 2.7), because EWT could explain a large part of the variance in nitrogen. These multiple regression models generated non-significant coefficients, which impeded their further utility in nitrogen estimation by coupling with inversion of physically based models. The reason of having non-significant coefficients could be the correlation between EWT and other leaf traits (LMA,  $CHL_{area}$ ). However, the idea of using combinations of leaf traits to enhance explaining the variance of nitrogen should not be denied, and need to be further investigated with larger datasets.

The combination of empirical and physical methods contributes to the moderate estimation of leaf nitrogen content in fresh leaves. The advantage of using a physically based model is the possibility of generating a large database of simulated spectra by varying the input parameters (le Maire et al. 2008), which may be applied to other sites. Though analysis were performed for area-based leaf nitrogen content and leaf traits, our technique are also applicable to mass-based parameters. The area-based leaf traits derived from PROSPECT model can be transformed to mass-based expressions by LMA or SLA, which can be used in empirical relationships established on mass-based leaf traits (Field and Mooney 1986; Lloyd et al. 2013). It is important to estimate leaf nitrogen content on a global scale because it is an essential biodiversity variable (Asner and Martin 2008a; Pereira et al. 2013; Wright et al. 2004), and global maps could be generated if hyperspectral satellites such as Sentinel (Malenovsky et al. 2012), EnMAP (Kaufmann et al. 2008), and HypSPIRI (<http://hyspiri.jpl.nasa.gov/>) become operational. Our method holds great value for exploring the potential of up-scaling these results to canopy level by coupling with a vegetation canopy reflectance model (Clevers and Kooistra 2012; Omari et al. 2013; Verhoef and Bach 2007) using air-borne or space-borne hyperspectral remote sensing.

## **2.5 Conclusions**

Leaf nitrogen content in fresh leaves has not so far been estimated by physically based methods (Jacquemoud et al. 1996), but here we demonstrate an indirect estimation of leaf nitrogen by using leaf traits forcing retrieved from a physically based model inversion. In this study, the area-based nitrogen correlations with

leaf traits were found to be higher than the mass-based correlations. Regression models were derived for area-based leaf nitrogen content using highly correlated leaf traits (LMA, EWT and  $CHL_{area}$ ) as independent variables. The empirical models and retrieved leaf traits were combined to estimate leaf nitrogen content. Our results indicated that EWT was retrieved with a higher accuracy than LMA or  $CHL_{area}$  through the inversion of PROSPECT. Area-based leaf nitrogen content was estimated more accurately by regression models using EWT as predictor than LMA or  $CHL_{area}$ . A linear regression model using EWT as a predictor provided the most accurate estimation of nitrogen. The combination of empirical and physically based models serves as a reliable method for estimating leaf nitrogen, although its transferability needs to be explored using other datasets. Further investigation is also needed to up-scale the study to canopy level, coupled with a canopy radiative transfer model using air-borne or space-borne hyperspectral remote sensing. Regional and global mapping of leaf nitrogen will further our understanding of the photosynthesis process, net primary productivity, and carbon dynamics (Evans 1989; Heimann and Reichstein 2008; Ollinger and Smith 2005; Reich et al. 2006). Our study provides practical techniques for estimating leaf nitrogen, which will be beneficial to the assessment of biodiversity and ecosystem services.

Leaf nitrogen content indirectly estimated by leaf traits from the PROSPECT model

## Chapter 3

### **Applicability of the PROSPECT Model for Estimating Protein and Cellulose + lignin in Fresh Leaves<sup>\*</sup>**

---

<sup>\*</sup> This chapter is based on: Wang, Z., Skidmore, A.K., Wang, T., Darvishzadeh, R., and Hearne J. (2015). Applicability of the PROSPECT model for estimating protein and cellulose + lignin in fresh leaves. *Remote Sensing of Environment*, 168, 205-218.

## **Abstract**

Hyperspectral remote sensing of leaf biochemicals is critical for understanding many biochemical processes. Leaf biochemical contents (e.g., protein, cellulose and lignin) in fresh and dry leaves have been quantified from hyperspectral data using empirical models. However, they cannot be retrieved for fresh leaves by inverting radiative transfer models. We demonstrated the applicability of PROSPECT leaf optical properties model in the separation of specific absorption coefficients for protein and cellulose + lignin following a newly proposed algorithm, and evaluated the feasibility in estimating leaf protein and cellulose + lignin content through model inversion. Assessment was performed across a large variety of plant species benefiting from the Leaf Optical Properties Experiment (LOPEX) dataset. To alleviate ill-posed problems, inversion was performed over different spectral subsets. The PROSPECT model with newly calibrated specific absorption coefficients was able to accurately reconstruct leaf reflectance and transmittance. Leaf protein and cellulose + lignin were estimated at moderate to good accuracies for both fresh and dry leaves. The spectral subset of 2100 - 2300 nm yielded the most accurate estimation of leaf cellulose + lignin ( $R^2 = 0.70$ , RMSE =  $5.21\text{E-}04$  g/cm<sup>2</sup>) and protein ( $R^2 = 0.47$ , RMSE =  $2.75\text{E-}04$  g/cm<sup>2</sup>) in fresh leaves, which were comparable with those obtained from stepwise multiple linear regressions (protein:  $R^2 = 0.83$ , RMSE= $3.91\text{E-}04$  g/cm<sup>2</sup>; cellulose + lignin:  $R^2 = 0.66$ , RMSE= $2.02\text{E-}04$  g/cm<sup>2</sup>). Our results confirm the importance of selecting a proper spectral subset that contains sufficient information for a successful inversion. For the first time, we provide promising estimations of leaf protein in fresh leaves through inversion of a radiative transfer model, which can be applied at canopy level for regional mapping if coupled with a canopy reflectance model and air- or space-borne hyperspectral imaging.

### 3.1 Introduction

Leaf protein, cellulose and lignin provide critical information about many biochemical processes, such as photosynthesis, respiration, and litter decomposition (Peterson and Hubbard 1992). Proteins are the main nitrogen-containing biochemical constituent in plants (Kokaly et al. 2009), and nitrogen is often a limiting factor for plant growth and important in terrestrial ecosystem carbon dynamics (Ollinger and Smith 2005; Smith et al. 2002). Cellulose and lignin (hereafter referred to as “cellulose + lignin”) is highly correlated with the total carbon in leaf (Jacquemoud et al. 1996), and play a significant role in forest litter decomposition and nutrient cycling (Aber and Federer 1992; Steudler et al. 1989). Therefore, remote estimation of leaf protein and cellulose + lignin are essential in producing an equivalent ratio of carbon/nitrogen (C/N), which will further improve the understanding of biogeochemical processes.

Two different techniques have been investigated for retrieving leaf protein and cellulose + lignin using hyperspectral data. Empirical methods have been the dominant techniques used to estimate the leaf parameters, focusing on building statistical regression models from the high correlations between their content and reflectance or its derivatives (Curran 1989; Kokaly and Clark 1999; Skidmore et al. 2010; Zhao et al. 2013). However, it proved difficult to transfer the empirical relationships across species or sites (LaCapra et al. 1996; Martin and Aber 1997). Moreover, the wavebands selected by statistical analysis from different studies have, on occasion, been inconsistent depending on using reflectance or transmittance, dry or fresh leaves, and have deviated from known absorption bands of leaf compounds (Curran et al. 2001; Huang et al. 2004; Jacquemoud et al. 1996).

Radiative transfer models (RTMs) describe the interaction of solar radiation in leaves based on laws of optics, thus offer advantages in robustness and transferability compared to empirical models (Darvishzadeh et al. 2011; Jacquemoud and Baret 1990; Schlerf and Atzberger 2006). RTMs have been widely applied to retrieve several leaf biochemical parameters (e.g., leaf chlorophyll, dry matter and water content) from remotely sensed data (Baret and Fourty 1997a; Darvishzadeh et al. 2008; Feret et al. 2008; Omari et al. 2013), but they have not been well developed for leaf protein and cellulose + lignin. Several studies (Fourty et al. 1996; Jacquemoud et al. 1996) attempted to incorporate leaf protein, cellulose + lignin into the absorption and scattering processes in radiative transfer models, such as the PROSPECT leaf optical properties model

(Jacquemoud and Baret 1990). Leaf protein and cellulose + lignin for dry leaves could be moderately well estimated ( $R^2 = 0.49 - 0.84$ ) through PROSPECT model inversion (Fourty et al. 1996; Jacquemoud et al. 1996), but so far, these parameters have not been successfully estimated for fresh leaves using the PROSPECT model (Botha et al. 2006; Jacquemoud et al. 1996; Kokaly et al. 2009). On fresh leaves, leaf reflectance and transmittance are insensitive to protein because of its small percentage in the leaf mass (Baret and Fourty 1997b; Jacquemoud et al. 1996). Moreover, the high covariance with water and other nitrogen-containing compounds, such as chlorophyll, has led to inconsistencies in retrieving leaf protein via PROSPECT inversion (Jacquemoud and Baret 1990; Jacquemoud et al. 1996; Kokaly et al. 2009). The idea of incorporating leaf protein, and cellulose + lignin into the PROSPECT model was therefore abandoned in the 1990s (Jacquemoud and Baret 1990; Jacquemoud et al. 1996), and they were represented by a more general parameter of “dry matter content” (Fourty et al. 1996).

The retrieval of leaf parameters (e.g. protein and cellulose + lignin) in fresh leaves by inversion of PROSPECT remains challenging. The first issue needed to be resolved is the determination of specific absorption coefficients of leaf biochemicals, which describes the absorption property of a leaf, together with the concentration of these biochemicals (Jacquemoud and Baret 1990). The specific absorption coefficients are considered to be inherent properties, which are constant across samples and species (Feret et al. 2008). It is difficult to determine these coefficients for the model; the coefficient for pure liquid water and chlorophyll have been well investigated (Buiteveld et al. 1994; Kou et al. 1993; Wieliczka et al. 1989), but gaps still exist for leaf pigments (i.e. chlorophyll a and b) and other cell wall constituents (i.e. protein, cellulose + lignin) (Feret et al. 2008). The obstacles lie in measuring the coefficients *in vivo*, as well as extracting each biochemical constituent for its measurement *in vitro* (Porra 2002). Therefore, the *in vivo* specific absorption coefficients are normally calibrated from intact leaves via model inversion using the measured spectra and biochemical contents (Feret et al. 2008; Fourty et al. 1996; Jacquemoud and Baret 1990; Jacquemoud et al. 1996).

The “ill-posed” inverse problem of radiative transfer model is the second challenge in the retrieval of leaf protein and cellulose + lignin. Different combinations of leaf parameters can generate the similar leaf spectra. And the leaf parameters which make greater contribution to leaf reflectance, such as leaf structure parameter and leaf water, bring more uncertainties in the retrieval of



protein and cellulose + lignin. Several studies have demonstrated that selecting spectral subsets will return a higher accuracy than using full wavelengths in inverting radiative transfer models (Darvishzadeh et al. 2008; Meroni et al. 2004; Schlerf and Atzberger 2006). The reasons for this can be poorly measured spectra, some wavelengths may not be well described in the model (Schlerf and Atzberger 2006), or extra bands may add noise without adding significant information (Weiss et al. 2000). In addition, higher accuracies could be obtained for estimating leaf parameters if the inversion approach would consider a specific merit function assigned for each parameter and incorporate sensitive spectral bands determined by sensitivity analysis (Li and Wang 2011; Zhao et al. 2014).

The radiative transfer model PROSPECT has been recalibrated with a newly proposed algorithm which can simulate leaf optical properties with a higher accuracy than previous versions (Feret et al. 2008). In their study, the incidence angle of incoming radiation was reassessed, and the specific absorption coefficients of leaf constituents and the refractive index were recalibrated for two new version models (PROSPECT-4 and PROSPECT-5). PROSPECT-5 is the same as PROSPECT-4 except the separation of total chlorophylls and total carotenoids in the visible range of 400 – 750 nm. The algorithm has been tested for accurate simulation of leaf reflectance (Li and Wang 2011; Ma et al. 2012), and has been further adapted to incorporate tannin as an input parameter for assessing the quality of tea in Tea-PROSPECT (Bian 2013). However, to our knowledge, the newly proposed algorithm has not been evaluated for determining the specific absorption coefficients for leaf constituents, such as protein, cellulose, hemicellulose, and lignin. And the potential of PROSPECT-5 for retrieving leaf constituents through model inversion needs to be further explored.

The estimation of leaf constituents (e.g. protein and cellulose + lignin) using fresh leaf spectra through the radiative transfer model PROSPECT has also not been studied extensively (Botha et al. 2006; Jacquemoud et al. 1996), probably because of the notion that there may be little sensitivity of leaf optical properties to the leaf constituents. The aim of our study was therefore to evaluate the most recent PROSPECT model (version 5) for estimating leaf protein, cellulose + lignin by recalibrating their specific absorption coefficients, using the calibration algorithm proposed by Feret et al. (2008). Our specific objectives were: (1) to test if the specific absorption coefficients could be robustly estimated across species from a model calibration procedure; (2) to test if the PROSPECT model with the recalibrated coefficients could successfully simulate leaf reflectance and transmittance in forward mode; (3) to evaluate if selected spectral subsets could

improve the estimation accuracy of leaf protein and cellulose + lignin, when compared to using the full wavelengths of 1100-2500 nm; and (4) to compare the accuracy of estimated leaf protein and cellulose + lignin from PROSPECT model inversion and stepwise multiple linear regression.

## **3.2 Materials and Methods**

### **3.2.1 Description of the dataset**

The Leaf Optical Properties Experiment (LOPEX) dataset was established by the Joint Research Center (JRC) of the European Commission (Ispra, Italy) in 1993. The dataset includes 50 woody and herbaceous species, representing a large variety of leaf internal structure, biochemical composition and spectra. Hemispherical reflectance and transmittance over 400-2500 nm were measured with a Perkin Elmer Lambda 19 Spectrophotometer equipped with an integrating sphere. Spectralon references were used for calibration. For each sample, reflectance and transmittance from five leaves were obtained, and averaged after calibration. The fresh weight (FW, g) of a 4.10 cm<sup>2</sup> (leaf area, LA) disk taken on each leaf using a cork borer was measured first, and reweighed for dry weight (DW, g) after the disk was placed in a drying oven at 85°C for 48 hours. Leaf mass per area (LMA, g/cm<sup>2</sup>) and equivalent water thickness (EWT, cm) were then calculated as:  $LMA = DW / LA$ , and  $EWT = (FW - DW) / LA$ . The biochemical constituents, including protein, cellulose, lignin and starch, were determined by two independent laboratories in France and Belgium. Standard wet chemistry analysis were performed for measuring protein using Kjeldahl method (AOAC 1970), cellulose (Van Soest and Wine 1991; Weende 1985), lignin (Van Soest and Masson 1967; Van Soest and Wine 1991) and starch (Ewerts 1985). Measurements from the two laboratories were quite consistent for protein and the sum of cellulose and lignin, and the data from the French laboratory were used.

In this study we only considered protein, cellulose and lignin, since starch is a minor constituent. Protein, cellulose and lignin were expressed as grams per unit fresh leaf area (g/cm<sup>2</sup>). In total, 63 fresh leaf samples and 49 dry leaf samples were studied. The statistical characteristics of these samples are presented in Table 3.1, and details of the dataset used and their measurements can be found in Hosgood et al. (1995) and Jacquemoud et al. (1996).

**Table 3.1** Statistics of the LOPEX dataset used in this study.

Parameter	Unit	Min	Max	Mean	Std.
Equivalent water thickness (EWT) <sup>a</sup>	cm	0.0046	0.0405	0.0114	0.0067
Equivalent water thickness (EWT) <sup>b</sup>	cm	1.85E-05	0.0009	0.0002	0.0001
Protein	g/cm <sup>2</sup>	2.92E-04	0.0020	0.0010	0.0003
Cellulose + lignin	g/cm <sup>2</sup>	1.47E-04	0.0046	0.0013	0.0010
Leaf mass per area (LMA) <sup>a</sup>	g/cm <sup>2</sup>	0.0019	0.0137	0.0053	0.0024
Leaf mass per area (LMA) <sup>b</sup>	g/cm <sup>2</sup>	0.0024	0.0165	0.0059	0.0026

<sup>a</sup> and <sup>b</sup> were for parameters in fresh and dry leaves, respectively.

### 3.2.2 Calibration of the PROSPECT model

The PROSPECT leaf optical properties model was developed to simulate leaf directional-hemispherical reflectance and transmittance over the optical domain from 400 nm to 2500 nm (Jacquemoud and Baret 1990). The PROSPECT model considers the leaf as a stack of elementary homogeneous layers, with each layer as an element with absorbing and scattering properties (Jacquemoud and Baret 1990). The scattering process is described by the number of stacked elementary homogeneous layers, leaf structure parameter  $N$  and the refractive index  $n(\lambda)$ . The absorption process is described by the absorption coefficient of one elementary layer  $k(\lambda)$ . The directional-hemispherical reflectance  $R_{mod}(\lambda)$  and transmittance  $T_{mod}(\lambda)$  at each wavelength  $\lambda$  can be simulated as:

$$[R_{mod}(\lambda), T_{mod}(\lambda)] = \text{PROSPECT}(N, k(\lambda), n(\lambda)) \quad (3.1)$$

where  $k(\lambda)$  is linear combination of the content of leaf constituents  $C_i$  (e.g., chlorophyll, water, and dry matter), and their corresponding specific absorption coefficients  $k_i(\lambda)$ :

$$k(\lambda) = \sum_i k_i(\lambda) \times \frac{C_i}{N} \quad (3.2)$$

Then Eq. (1) can be simplified as follows by using only four input parameters:

$$[R_{mod}(\lambda), T_{mod}(\lambda)] = \text{PROSPECT}(N, \text{Chl}, \text{LMA}, \text{EWT}) \quad (3.3)$$

where the refractive index and specific absorption coefficients are fixed as constant values from the model calibration. The parameters are leaf chlorophyll content (Chl,  $\mu\text{g}/\text{cm}^2$ ), LMA ( $\text{g}/\text{cm}^2$ ), and EWT (cm), respectively. Note that Eq. (1) differs from Eq. (3.3), because the influences of Chl, LMA and EWT on leaf optical properties were represented by the parameter  $k(\lambda)$ , using Eq. (3.2).

The improved (1-nm resolution) and recalibrated version, PROSPECT-5, which was published in 2008 (Feret et al. 2008), is used in this study. We performed the recalibration in three steps following the algorithm presented in Feret et al. (2008). The first step was to determine the leaf structure parameter through inversion of the PROSPECT model using the “three wavelengths method” (Feret et al. 2008; Jacquemoud et al. 1996). The maximum incident angle  $\alpha$  was set to a more realistic value of  $40^\circ$  in this study, as suggested by PROSPECT-4 (Feret et al. 2008). In the second step, we calculated the absorption coefficient of one elementary layer  $k(\lambda)$  and the refractive index  $n(\lambda)$  by inverting the model for a second time and using the leaf structure parameter determined in the first step. The derived refractive index in this study was close to that in PROSPECT-5, and therefore the refractive index from PROSPECT-5 was adopted for later analysis, as that calibration involved more comprehensive datasets (Feret et al. 2008).

In the third step, we deduced the specific absorption coefficients of leaf protein and cellulose + lignin from Eq. (3.2), using the absorption coefficient of one elementary layer  $k(\lambda)$  computed in the second step, together with the measured leaf water, protein, cellulose + lignin content from the LOPEX dataset. As suggested by Feret et al. (2008), the specific absorption coefficient of water did not need recalibration, so we used the values in PROSPECT-5. In this step, we divided the dry matter into protein, cellulose and lignin, and chose to study the 800–2500 nm region, because water, protein, cellulose and lignin absorbs solar radiation mostly in this range (Baret and Fourty 1997a; Curran 1989; Feret et al. 2008). We combined the cellulose and lignin as one constituent (“cellulose + lignin”). Different combinations of leaf constituents were tested as input parameters to PROSPECT by Jacquemoud et al. (1996), who found that the combination of water, protein, and cellulose + lignin best explained the variation in leaf optical properties in the 800–2500 nm range (Jacquemoud, personal communication).

The inversion was performed using a bounded optimization package, FMINSEARCHBND.M in MATLAB (The MathWorks, Inc.). The MATLAB code is available online provided by D'Errico (2012). The solution method only guarantees a local minimum, but after extensive tests with sets of initial values it was concluded that a global minimum had been found. The calibration process was performed separately for fresh, dry and fresh + dry leaves. The recalibrated specific absorption coefficients of protein, and cellulose + lignin from each group were compared with previous results reported by Jacquemoud et al. (1996).

### 3.2.3 Sensitivity analysis

Sensitivity analysis (SA) helps to identify the contribution of the variability in input parameters to the variability in the output reflectance or transmittance. Two types of SA have been listed in literature, local SA and global SA. Local SA provides information on how the variation of each input parameter individually explains the variation in the model output and ignores the interactions between model parameters (Asner 1998; Saltelli 1999), while global SA provides information of how the variation of each input parameter individually and their interactions with each other accounts for the variation of model output. Global SA is required when searching the full range of model parameter space, including the simultaneous variations of model parameters (Bowyer and Danson 2004), therefore, a global SA was adopted in this study.

SA calculates the fractional contribution of a given input variable  $X$  to the variance of  $Y$ . Supposing a model with three input parameters (Ceccato et al. 2002), then the total variance  $V(Y)$  of the model output is decomposed as followed

$$V = V_1 + V_2 + V_3 + V_{12} + V_{13} + V_{23} + V_{123} \quad (3.4)$$

where  $V_1$  is the variance of input parameter 1,  $V_{12}$  is the variance of interaction between parameters 1 and 2, and  $V_{123}$  is the variance of interaction between parameters 1, 2 and 3, and so on.

The first-order sensitivity index  $S_1$  for parameter 1 is defined as

$$S_1 = V_1 / V \quad (3.5)$$

Similarly, the second-order sensitivity index  $S_{12}$  can be calculated from

$$S_{12} = V_{12} / V \quad (3.6)$$

And the total-order indices  $ST$  for parameter 1 is measured as follows

$$ST = (V_1 + V_{12} + V_{13} + V_{123}) / V \quad (3.7)$$

The first-order sensitivity indices represents the independent effect of the corresponding parameters while the second and higher order terms gives the interaction effects of each parameter with others. The Fourier amplitude sensitivity test (FAST) method calculates the contribution of each input parameter to the variation in the output efficiently (Saltelli and Bolado 1998; Xiao et al. 2014), which refers to first-order sensitivity indices. The method of Sobol (Sobol 1993) was superior to FAST in computing the higher interaction terms

(DeJonge et al. 2012; Xiao et al. 2014). A Matlab software tool (GSAT) (Cannavó 2012) was applied in this study to perform global sensitivity analysis, in which both FAST analysis and Sobol's analysis were implemented to calculate first order sensitivity coefficients and joint global sensitivity coefficients for model parameter combinations. The global sensitivity analysis was separately performed for fresh and dry leaves using the ranges of input parameters presented in Table 3.1.

#### **3.2.4 Validation of the recalibrated PROSPECT-5 model in forward mode**

The k-fold cross validation was performed due to the limited number of samples, which enables independent data set used in the calibration and validation. The commonly used 10-fold cross-validation was adopted in this study. Detailed descriptions are presented as follows.

- (1) The original samples were randomly partitioned into k equal size subsamples (k=10).
- (2) The k-1 subsamples were used in the calibration to determine the specific absorption coefficients of protein, and cellulose + lignin (see Section 2.2).
- (3) The remaining one subsample was retained as the validation data for testing the performance of the recalibrated model in forward modeling leaf reflectance and transmittance, by using the calibrated specific absorption coefficients from the previous step, the measured content of constituents (e.g., water, protein, cellulose + lignin) and estimated leaf structure parameter in the calibration.
- (4) The cross-validation procedure was then repeated k times with each one of the k subsamples used once as the validation data, after which the modelled leaf reflectance and transmittance of all samples can be obtained.

The k-fold cross validation (step 1-4) was repeated 100 times to avoid biased results (Schlerf et al. 2005). We used the measured content of constituents (e.g., water, protein, cellulose + lignin), their corresponding recalibrated specific absorption coefficients, and the estimated leaf structure parameter to simulate the reflectance and transmittance in the forward mode. Simulated leaf reflectance and transmittance were also obtained by using the specific absorption coefficients presented by Jacquemoud et al. (1996). For comparison, this simulation was also implemented using the measured content of constituents (e.g., water, dry matter), their specific absorption coefficients in PROSPECT-5, and the estimated leaf structure parameter.

The root mean square error of prediction (RMSEP), the bias (BIAS), and the standard error of prediction corrected from the bias (SEPC) between the modeled and measured reflectance and transmittance were calculated for each wavelength. The averaged RMSEP, BIAS, SEPC of the 100 repetitions were used to evaluate the model performance using calibrated specific absorption coefficients from this study. Let  $y_j$  and  $y_j'$  be the measured and estimated values, respectively,  $\bar{y}$  be the mean of the measured values, and  $n$  be the number of measurements:

$$\text{RMSEP} = \sqrt{\frac{\sum_{j=1}^n (y_j' - y_j)^2}{n}} \quad (3.8)$$

$$\text{BIAS} = \frac{\sum_{j=1}^n (y_j' - y_j)}{n} \quad (3.9)$$

$$\text{SEPC} = \sqrt{\frac{\sum_{j=1}^n (y_j' - y_j - \text{BIAS})^2}{n}}. \quad (3.10)$$

### 3.2.5 Estimation of leaf parameters

#### 3.2.5.1 PROSPECT model inversion

The input model parameters included leaf structure parameter ( $N$ ), EWT, protein content ( $C_p$ ), and cellulose + lignin content ( $C_{cl}$ ). The inversion process was to find the parameter vector  $\theta = [N, \text{EWT}, C_p, C_{cl}]^T$ , which minimizes the merit function:

$$J(\theta) = \sum_{\lambda_1}^{\lambda_2} (R_{mes}(\lambda) - R_{mod}(\lambda, \theta))^2 + (T_{mes}(\lambda) - T_{mod}(\lambda, \theta))^2 \quad (3.11)$$

where  $\lambda$  is the wavelength,  $R_{mes}$  and  $T_{mes}$  are respectively the measured reflectance and transmittance, and  $R_{mod}$  and  $T_{mod}$  are the modeled values. The ranges of the input parameters were determined based on the dataset used in this study (Table 3.1) and are given in Table 3.2. The parameters were estimated through model inversion using an iterative optimization algorithm. Specific absorption coefficients of protein, and cellulose + lignin determined by the pooled samples were utilized to involve the variations of leaf optical and biochemical properties across species.

**Table 3.2** Ranges of the input parameters used in the recalibrated PROSPECT-5 model.

<b>Parameter</b>	<b>Unit</b>	<b>Min</b>	<b>Max</b>
<b>Leaf structure parameter (<i>N</i>)</b>	-	1	3
<b>Equivalent water thickness (EWT) <sup>a</sup></b>	cm	0.001	0.05
<b>Equivalent water thickness (EWT) <sup>b</sup></b>	cm	0.00001	0.001
<b>Protein</b>	g/cm <sup>2</sup>	0.0001	0.002
<b>Cellulose + lignin</b>	g/cm <sup>2</sup>	0.0001	0.005

<sup>a</sup> and <sup>b</sup> were for parameters in fresh and dry leaves, respectively.

To alleviate the “ill-posed” inverse problems, a specific merit function was assigned for each parameter defined on its own sensitive spectral bands, which has been proved to improve the performance of inversion results (Li and Wang 2011; Wang et al. 2015b; Zhao et al. 2014). Thus, we performed the model inversion for retrieving leaf parameters not only over the full wavelength range (1100-2500 nm, subset 1), but also over another six spectral subsets (Table 3.3). Subset 2 was the sensitive wavelength region for protein and cellulose + lignin, according to the global sensitivity analysis for fresh leaves (see Section 2.3). Subset 3 was also defined for protein and cellulose + lignin, but based on the global SA for dry leaves. Subset 4 was a collection of protein absorption bands from previous findings (Curran 1989; Fourty et al. 1996). Subset 5 was a collection of cellulose + lignin absorption bands (Curran 1989; Fourty et al. 1996). Subset 4 and 5 were involved to assess the plausibility of the calibrated specific absorption coefficients of protein and cellulose + lignin, and to test whether the subsets could achieve an equal or even improved estimation accuracy of leaf protein and cellulose + lignin compared to using the full wavelengths.



**Table 3.3** Selected spectral subsets for model inversion to estimate leaf protein and cellulose + lignin content based on absorption features from the literature, global sensitivity analysis and stepwise multiple linear regressions.

<b>Spectral subset</b>	<b>Wavelengths (nm)</b>
<b>1</b>	1100-2500
<b>2</b>	1600-1800, 2100-2300
<b>3</b>	2100-2300
<b>4</b>	1020, 1510, 1730, 1980, 2060, 2130, 2180, 2240, 2300
<b>5</b>	1120, 1200, 1420, 1450, 1490, 1540, 1780, 1820, 2100, 2270, 2280, 2340
<b>6<sup>a</sup></b>	1345, 2120, 2295, 2300
<b>6<sup>b</sup></b>	1320, 1450, 2015, 2290
<b>7<sup>a</sup></b>	1150, 1185, 1220, 1295
<b>7<sup>b</sup></b>	2000, 2150, 2170, 2255

Spectral subset 2 and 3 are selected based on global sensitivity analysis. 4 and 5 are associated with protein absorption features and cellulose + lignin absorption features, respectively. See (Curran 1989; Fourty et al. 1996). 6<sup>a</sup> and 7<sup>a</sup> are selected spectral subsets for protein and cellulose + lignin, respectively, based on stepwise multiple linear regressions on fresh leaves. 6<sup>b</sup> and 7<sup>b</sup> are selected spectral subsets for protein and cellulose + lignin, respectively, based on stepwise multiple linear regressions on dry leaves.

### 3.2.5.2 Stepwise multiple linear regression

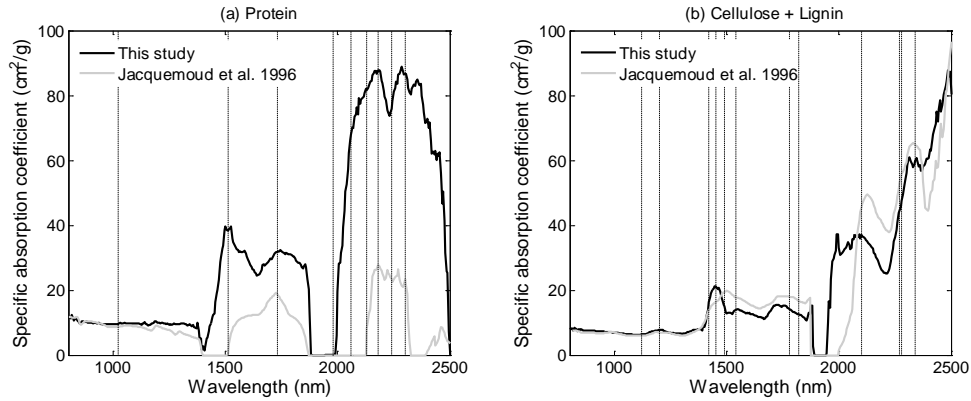
We also applied a classical statistical technique, stepwise multiple linear regression (SMLR) to select sensitive spectral subsets for protein and cellulose + lignin, and to compare its performance of estimation with that of PROSPECT model. Though a number of limitations have been pointed out for SMLR (Majeke et al. 2008), it serves as a straightforward way for comparison with inversion models. The stepwise multiple linear regression relates the protein or cellulose + lignin content with the reflectance or transmittance. The model starts with no independent variables (e.g. reflectance), adds the variable which was the best statistical predictor, and adds the second variable which gives the best fit together with the first variable. Further variable was added to provide an optimal fit in an recursive way (Jacquemoud et al. 1995). The SMLR was performed using MATLAB (The MathWorks, Inc.) over the 1100-2500nm in order to compare results with those from PROSPECT inversion. We restricted the number of selected wavelengths to four to have a simple model and avoid over-fitting. Subset 6 and 7 in Table 3.3 were determined on the selected wavelengths by SMLR for predicting protein and cellulose + lignin, respectively.

The accuracies of the retrieved parameters (EWT,  $C_p$ ,  $C_{cl}$ ) via PROSPECT model inversion and SMLR were evaluated by the coefficient of determination ( $R^2$ ), the root mean square error (RMSE), and the normalized RMSE (NRMSE = RMSE/range) between the estimated and measured values.

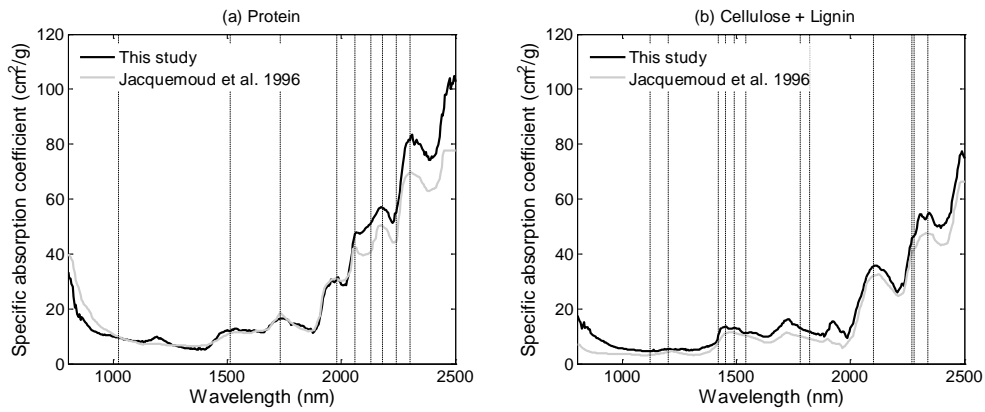
### **3.3 Results**

#### **3.3.1 Recalibrated specific absorption coefficients for protein and cellulose + lignin**

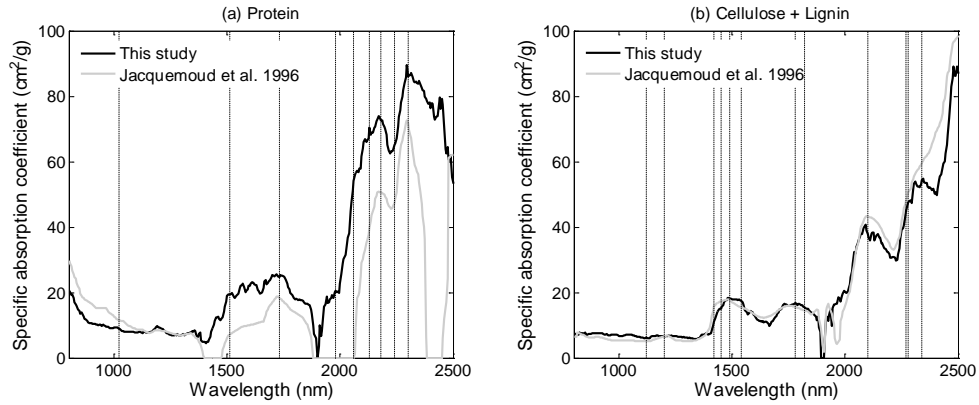
The recalibrated specific absorption coefficients for protein and cellulose + lignin over 800-2500 nm determined on fresh leaves, dry leaves, and fresh + dry leaves are shown in Fig. 3.1, Fig. 3.2, and Fig. 3.3, respectively. For dry leaves, most of the absorption peaks of protein and cellulose + lignin were well represented (Fig. 3.2), and are consistent with findings by Jacquemoud et al. (1996). For fresh leaves, fewer absorption peaks were displayed (see dashed vertical lines in Fig. 3.1), and discrepancies were found between the absorption peaks in specific absorption coefficients and absorption features in literature (Table 3.4 and 3.5). Generally, absorption positions were consistent for coefficients determined on fresh leaves, dry leaves and fresh + dry leaves, except for those masked by the strong water absorption at 1800-2000 nm. 100 repetitions of k-fold cross-validation generated quite consistent specific absorption coefficients for protein and cellulose + lignin (Fig. 3.4 and 3.5). The coefficients of protein and cellulose + lignin for dry leaves display minor differences with results in a previous study (Jacquemoud et al. 1996). However, the recalibrated specific absorption coefficient of protein for fresh leaves yielded significantly different results. The coefficients of protein for fresh leaves in this study are much higher (with a peak value of over 80 cm<sup>2</sup>/g) than results reported by Jacquemoud et al. (1996).



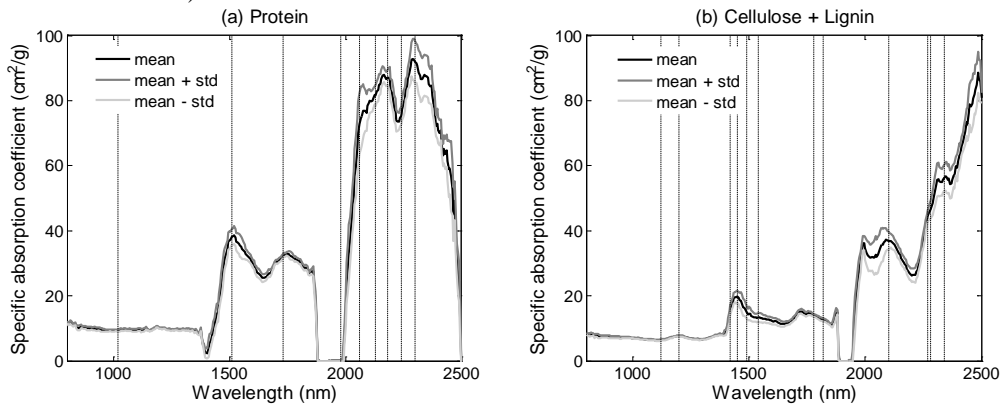
**Figure 3.1** Recalibrated specific absorption coefficients of protein (a) and cellulose + lignin (b) determined on fresh leaves. The vertical dashed lines correspond to published absorption features of protein (a) and cellulose + lignin (b) (see details in Table 3.4 and 3.5).



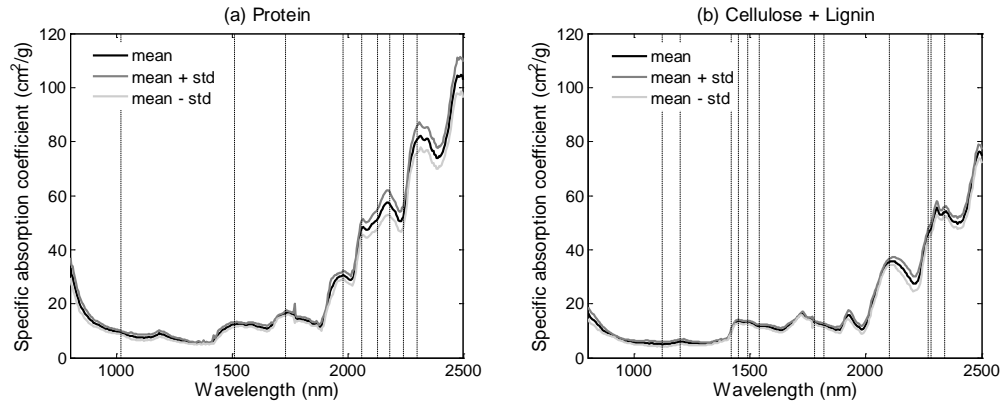
**Figure 3.2** Recalibrated specific absorption coefficients of protein (a) and cellulose + lignin (b) determined on dry leaves. The vertical dashed lines correspond to published absorption features of protein (a) and cellulose + lignin (b) (see details in Table 3.4 and 3.5).



**Figure 3.3** Recalibrated specific absorption coefficients of protein (a) and cellulose + lignin (b) determined on fresh + dry leaves. The vertical dashed lines correspond to published absorption features of protein (a) and cellulose + lignin (b) (see details in Table 3.4 and 3.5).



**Figure 3.4** Mean and variability of the recalibrated specific absorption coefficients of protein and cellulose + lignin from 100 repetitions of k-fold cross-validation (fresh leaves).



**Figure 3.5** Mean and variability of the recalibrated specific absorption coefficients of protein and cellulose + lignin from 100 repetitions of k-fold cross-validation (dry leaves).

**Table 3.4** Absorption peaks of protein reported in Curran (1989) and Fourty (1996), and in recalibrated specific absorption coefficients (SAC) from fresh, dry and fresh + dry leaves.

Wavelength (nm) Literature	Wavelength (nm) SAC from fresh leaves	Wavelength (nm) SAC from Dry leaves	Wavelength (nm) SAC from fresh + dry leaves	Absorption mechanisms
1020	-	-	-	N-H stretch
1510	1520	1510	1520	N-H stretch, 1st overtone
1730	1750	1730	1730	C-H stretch
1980	-	1980	-	N=H asymmetry
2060	-	2060	-	N=H bend, 2nd overtone, N=H bend, N-H stretch
2130	-	-	-	N-H stretch
2180	2165	2180	2170	N-H bend, 2nd overtone, C-H stretch, C-O stretch, C=O stretch C-N stretch
2240	-	-	-	C-H stretch
2300	2295	2300	2300	N-H bend, C=O stretch, C-H bend, 2nd overtone

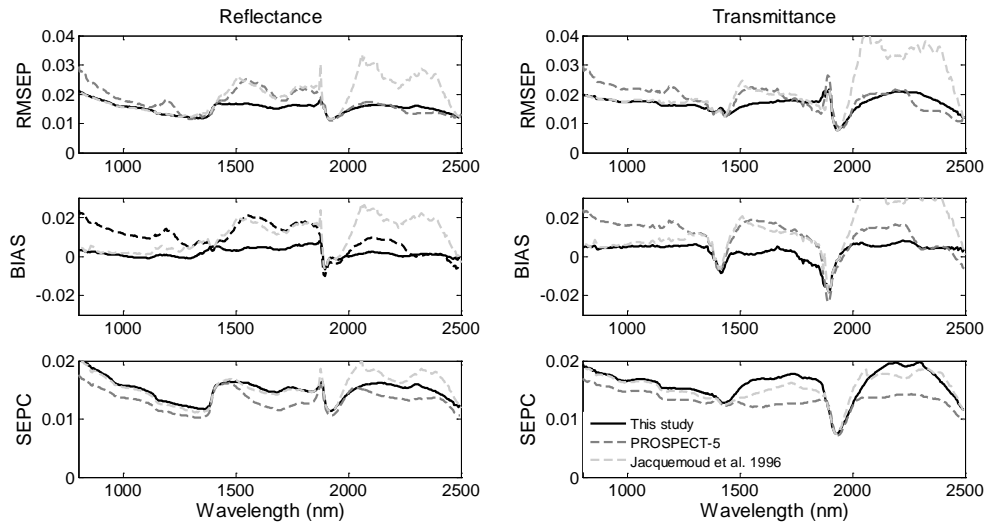
**Table 3.5** Absorption peaks of cellulose + lignin reported in Curran (1989) and Fourty (1996), and in recalibrated specific absorption coefficients (SAC) from fresh, dry and fresh + dry leaves.

Wavelength (nm) Literature	Wavelength (nm) SAC from fresh leaves	Wavelength (nm) SAC from Dry leaves	Wavelength (nm) SAC from fresh + dry leaves	Absorption mechanisms
1120	-	-	-	C-H stretch, 2nd overtone
1200	1200	1200	1200	O-H bend, 1st overtone
1420	-	-	-	C-H stretch, C-H deformation
1450	1450	1450	1460	O-H stretch, 1st overtone, C-H stretch, C-H deformation
1490	-	1490	1470	O-H stretch, 1st overtone
1540	-	1540	-	O-H stretch, 1st overtone
1780	1760	-	-	C-H stretch, 1st overtone, O-H stretch, H-O-H deformation
1820	-	-	-	O-H stretch, C-O stretch, 2nd overtone
2100	2100	2100	2100	O=H bend, C-O stretch, C- O-C stretch, 3rd overtone
2270	-	-	-	C-H stretch, O-H stretch, CH <sub>2</sub> bend, CH <sub>2</sub> stretch
2280	-	-	-	C-H stretch, CH <sub>2</sub> deformation
2340	2340	2340	2340	C-H stretch, O-H deformation, C-H deformation, O-H stretch

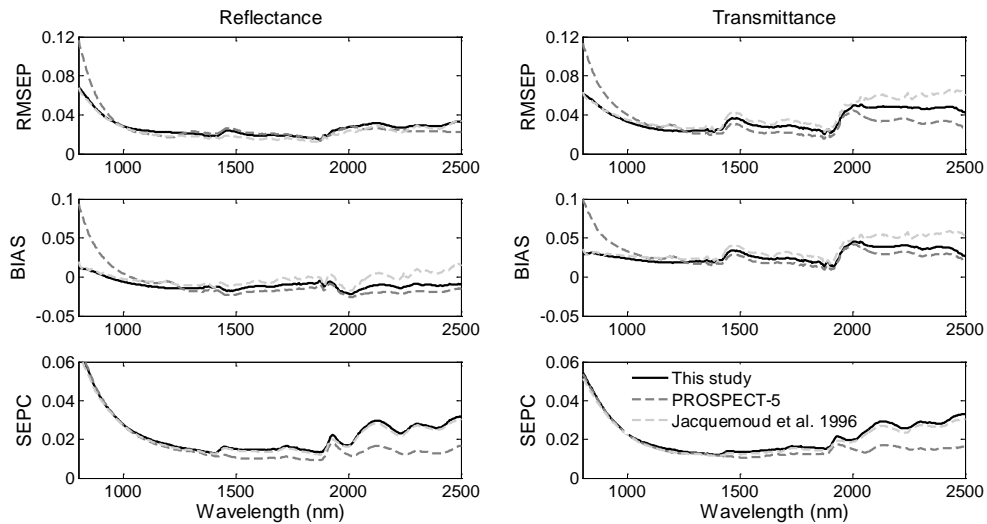
### 3.3.2 Validation of the performance of recalibrated PROSPECT model in forward mode

The modeling of leaf reflectance and transmittance as performed in this study is clearly more accurate than that with the original PROSPECT-5 (Fig. 3.6 and Fig. 3.7), despite the fact that the sum of protein and cellulose + lignin occupied two-thirds to three-quarters of the dry leaf mass, which was also confirmed by Jacquemoud et al. (1996). The spectral RMSEP's were generally less than 0.03, which confirms the capability of the PROSPECT model when recalibrated for simulating leaf optical properties using only a few parameters (Jacquemoud and Baret 1990; Jacquemoud et al. 1996). For fresh leaves (Fig. 3.6), the recalibrated

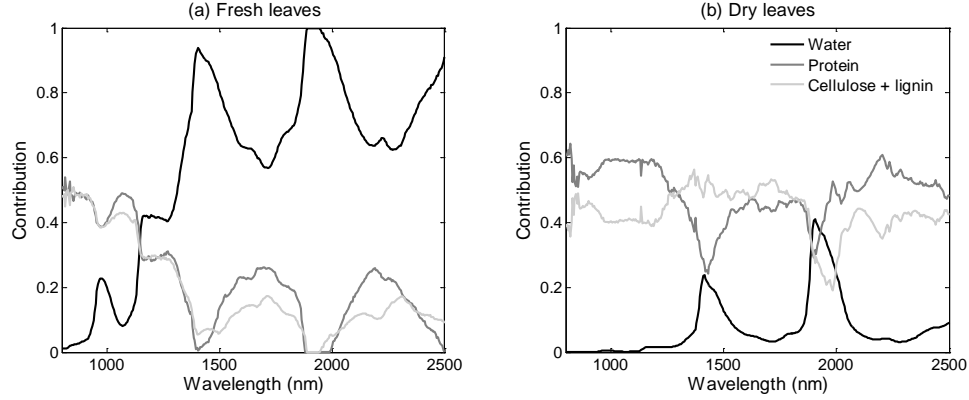
specific absorption coefficients greatly improved the performance of modeling leaf optical properties over the range of 1400-1800 nm and 2100-2500 nm. The modeled fresh leaf spectra using coefficients from Jacquemoud et al. (1996) gave the least accurate estimate in the range of 2100-2500 nm.



**Figure 3.6** RMSEP, BIAS and SEPC of the modeling of the reflectance and transmittance of fresh leaves by this study, Jacquemoud et al. (1996) and PROSPECT-5. RMSEP, BIAS and SEPC of this study were averaged values from 100 repetitions of k-fold cross validation.



**Figure 3.7** RMSEP, BIAS and SEPC of the modeling of the reflectance and transmittance of dry leaves by this study, Jacquemoud et al. (1996) and PROSPECT-5. RMSEP, BIAS and SEPC of this study were averaged values from 100 repetitions of k-fold cross validation.



**Figure 3.8** Contribution of water, protein and cellulose + lignin to the total leaf absorption (800-2500 nm) in fresh (a) and dry (b) leaves.

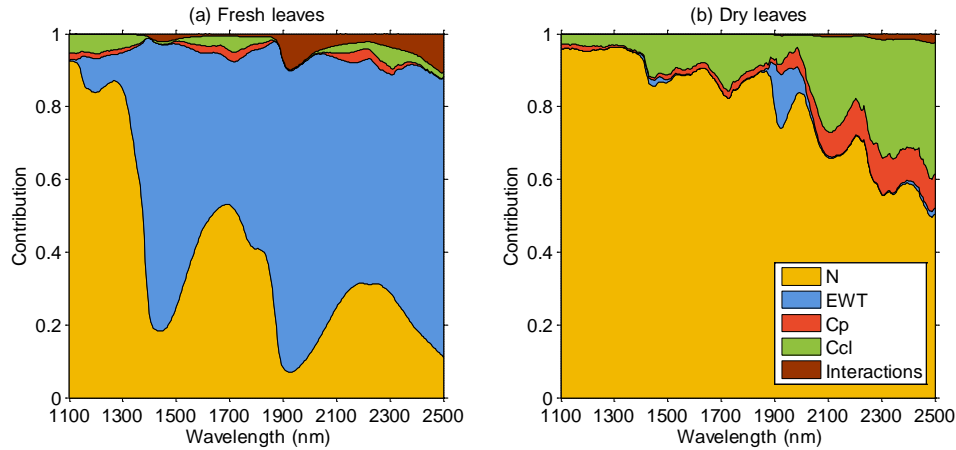
### 3.3.3 Contribution of different constituents to the total leaf absorption

The contribution of each constituent to the total leaf absorption coefficient was computed as (Baret and Fourty 1997a; Fourty et al. 1996):

$$k_i(\lambda) \cdot \bar{C}_i / (\sum_i k_i(\lambda) \times \bar{C}_i) \quad (3.12)$$

where  $\bar{C}_i$  is the average content of the constituent  $i$ , and  $k_i(\lambda)$  the corresponding specific absorption coefficient. This can provide information on identifying the sensitive bands for each constituent used in model inversion. Fig. 8 demonstrates the contribution of water, protein and cellulose + lignin to the total leaf absorption coefficient. It was not surprising to see the largest contribution of water content in fresh leaves over the whole range, and two peaks were located around 1400-1600 nm and 1800-2000 nm. Protein and cellulose + lignin have a relatively larger contribution over 1600-1800 nm and 2100-2300 nm. The contribution of protein was generally larger than that of cellulose + lignin. The protein and cellulose + lignin contributed approximately 40% to the total leaf absorption in these spectral ranges. For dry leaves, the contribution of water showed similar trends (in terms of spectral regions), but was generally lower (less than 50% contribution to total leaf absorption) than that of the other two constituents. Cellulose + lignin contributed most to leaf absorption over 1300-1800 nm, while protein contributed more after 1900 nm, which is consistent with reports by Fourty et al. (1996).





**Figure 3.9** Results of FAST first order sensitivity coefficients and interactions to leaf reflectance for global sensitivity analysis with PROSPECT model in (a) fresh leaves, (b) dry leaves. The legend is given in (b), where  $N$  is the leaf structure parameter, EWT is the equivalent water thickness,  $C_p$  is the protein content,  $C_{cl}$  is the cellulose + lignin content.

### 3.3.4 Global sensitivity analysis

Fig. 3.9 demonstrates the results of FAST first order sensitivity coefficients and interactions to leaf reflectance for global sensitivity analysis with PROSPECT model. Global sensitivity analysis showed that the sum of FAST first order sensitivity coefficients ( $N$ , EWT,  $C_p$ ,  $C_{cl}$ ) was close to 100% for both fresh and dry leaves, which indicated that a minor contribution was from interactions of these parameters to the leaf reflectance variability. For fresh leaves, leaf structure parameter  $N$  had the greatest contribution, followed by EWT. In contrast, the influence of cellulose + lignin and protein was limited, and was relatively greater in the spectral ranges of 800 - 1300 nm, 1600 - 1800 nm and 2100 - 2300 nm. Compared to fresh leaves, the influence of both protein and cellulose + lignin on the leaf spectra for dry leaves was much greater in wavelengths longer than 2100 nm, and the sum of their contributions amounted to around 40% (Fig. 3.9 (b)). Leaf structure parameter still acted as the dominant role in driving the leaf reflectance variability, but the influence of water was only noticeable in 1880 - 2000 nm. Similar features were observed for transmittance in global sensitivity analysis (results not shown).

### 3.3.5 Accuracy of estimated leaf parameters

The accuracy of the estimation for EWT, protein and cellulose + lignin content using the recalibrated PROSPECT-5 inversion is shown in Table 3.6 (fresh leaves) and Table 3.7 (dry leaves). For fresh leaves, EWT was estimated well by all seven

spectral subsets, with an  $R^2$  around 0.90, and RMSE of 0.0020 cm. Spectral subset 3 gave the most accurate estimation of cellulose + lignin ( $R^2 = 0.70$ ; RMSE =  $5.21\text{E-}04$  g/cm<sup>2</sup>), followed by subset 7<sup>a</sup> ( $R^2 = 0.61$ ; RMSE =  $8.48\text{E-}04$  g/cm<sup>2</sup>). The estimation of protein based on spectral subset 5 yielded the highest  $R^2$  value of 0.48 and a RMSE of  $3.12\text{E-}04$  g/cm<sup>2</sup>, followed by that based on the subset 3 ( $R^2 = 0.47$ ; RMSE =  $2.75\text{E-}04$  g/cm<sup>2</sup>). For dry leaves, EWT was generally not well assessed using all the spectral subsets, and cellulose + lignin values were estimated with a higher accuracy than protein values. Subset 3 with the region of 2100 – 2300 nm provided the most accurate retrieval for both cellulose + lignin ( $R^2 = 0.79$ ; RMSE =  $5.05\text{E-}04$  g/cm<sup>2</sup>) and protein ( $R^2 = 0.58$ ; RMSE =  $3.55\text{E-}04$  g/cm<sup>2</sup>). The subset 7 based on selected wavelengths for predicting cellulose + lignin using SMLR, gave a robust estimation for both fresh and dry leaves when applying the subset in model inversion. The selected subset 6<sup>b</sup> for protein based on SMLR on dry leaves also provide robust estimation in model inversion, but it was not valid for the subset 6<sup>b</sup> on fresh leaves. Table 3.8 presents the accuracy of predicted protein and cellulose + lignin using stepwise multiple linear regression. Generally, the stepwise multiple linear regression outperformed in estimations of protein and cellulose + lignin over PROSPECT inversion, especially for protein in fresh leaves (Fig. 3.10 and 3.11).

**Table 3.6** Validation of equivalent water thickness (EWT), protein and cellulose + lignin content retrieval for fresh leaves using recalibrated PROSPECT-5.

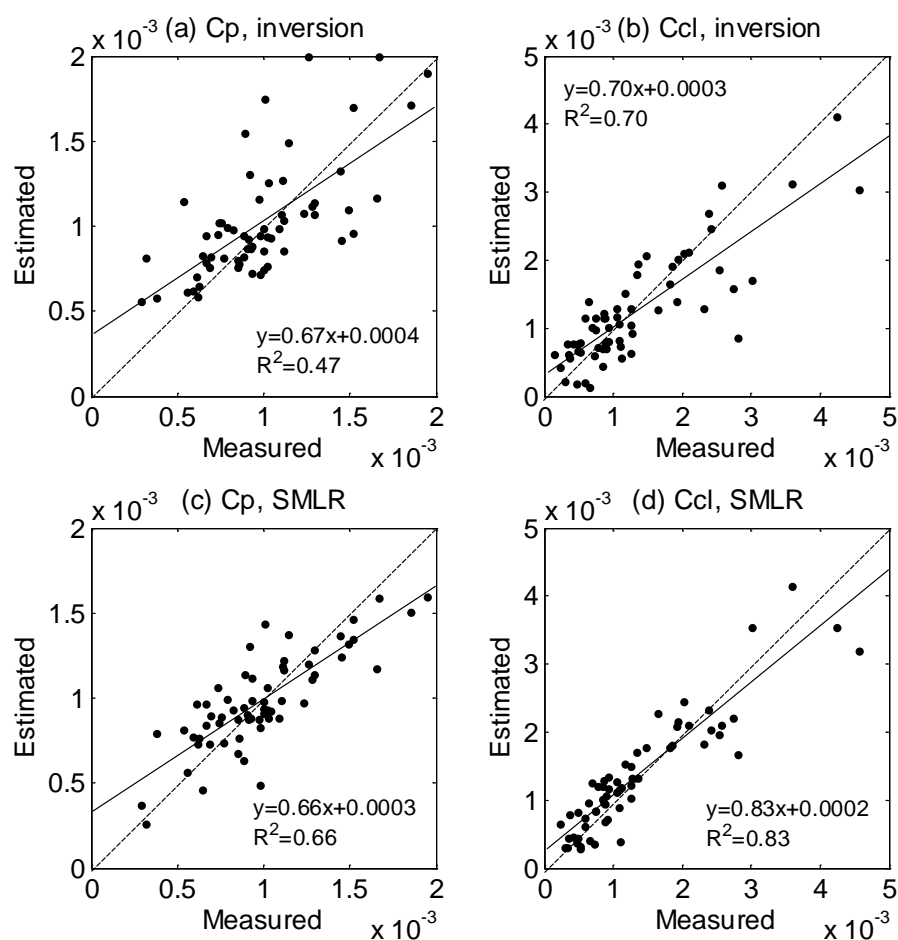
Spectral subset	EWT (cm)			Protein (g/cm <sup>2</sup> )			Cellulose + lignin (g/cm <sup>2</sup> )		
	$R^2$	RMSE	NRMSE	$R^2$	RMSE	NRMSE	$R^2$	RMSE	NRMSE
Subset 1	0.95	0.0015	0.04	0.45	$3.16\text{E-}04$	0.19	0.40	$8.66\text{E-}04$	0.20
Subset 2	0.89	0.0022	0.06	0.26	$3.81\text{E-}04$	0.23	0.44	$9.19\text{E-}04$	0.21
Subset 3	0.91	0.0020	0.06	0.47	$2.75\text{E-}04$	0.17	0.70	$5.21\text{E-}04$	0.12
Subset 4	0.94	0.0017	0.05	0.44	$3.11\text{E-}04$	0.19	0.46	$7.77\text{E-}04$	0.18
Subset 5	0.96	0.0014	0.04	0.48	$3.12\text{E-}04$	0.19	0.42	$8.24\text{E-}04$	0.19
Subset 6 <sup>a</sup>	0.83	0.0030	0.08	-	$8.13\text{E-}04$	0.49	0.33	$9.51\text{E-}04$	0.22
Subset 7 <sup>a</sup>	0.91	0.0029	0.08	-	$8.68\text{E-}04$	0.52	0.61	$8.48\text{E-}04$	0.19

**Table 3.7** Validation of equivalent water thickness (EWT), protein and cellulose + lignin content retrieval for dry leaves using recalibrated PROSPECT-4.

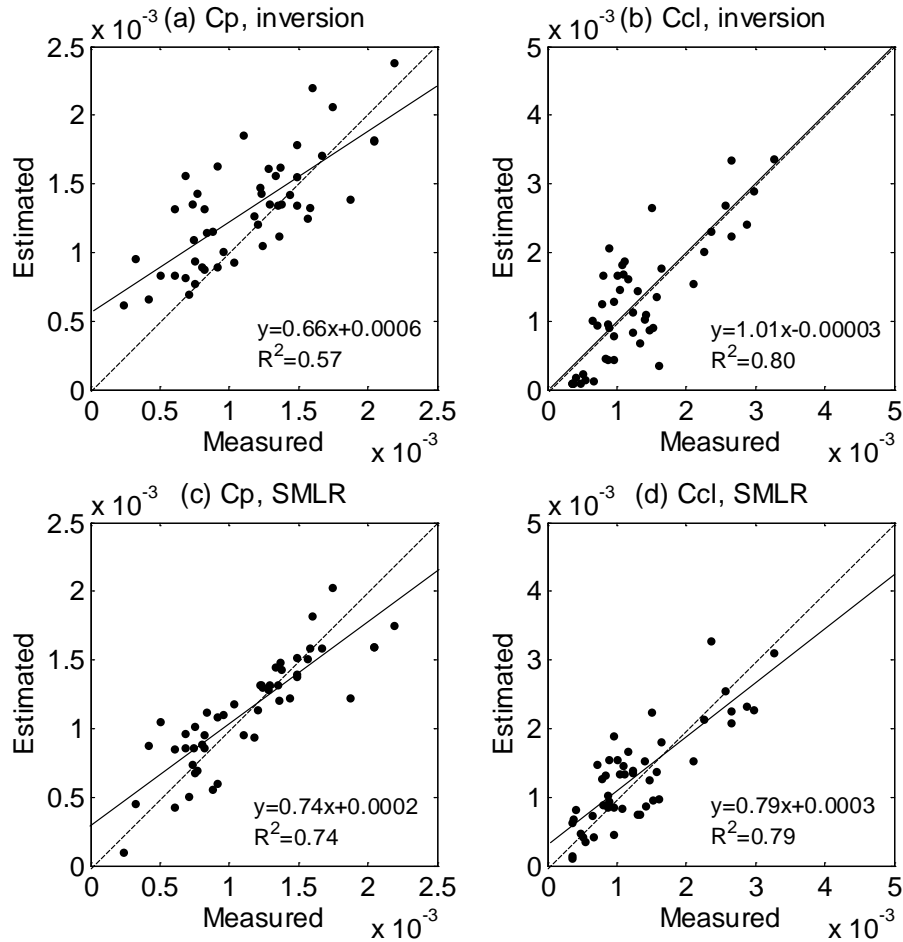
Spectral subset	EWT (cm)			Protein (g/cm <sup>2</sup> )			Cellulose + lignin (g/cm <sup>2</sup> )		
	$R^2$	RMSE	NRMSE	$R^2$	RMSE	NRMSE	$R^2$	RMSE	NRMSE
Subset 1	0.51	9.67E-05	0.11	0.50	3.96E-04	0.20	0.59	6.81E-04	0.13
Subset 2	-	4.96E-04	0.56	0.37	4.83E-04	0.25	0.41	8.24E-04	0.15
Subset 3	-	5.16E-04	0.58	0.57	3.55E-04	0.18	0.79	5.05E-04	0.09
Subset 4	0.32	1.06E-04	0.14	0.52	3.90E-04	0.20	0.68	5.72E-04	0.11
Subset 5	0.27	1.55E-04	0.18	0.43	4.84E-04	0.25	0.21	1.11E-03	0.21
Subset 6 <sup>b</sup>	0.32	1.43E-04	0.16	0.57	3.71E-04	0.19	0.62	6.88E-04	0.13
Subset 7 <sup>b</sup>	0.49	1.18E-04	0.13	0.11	7.61E-04	0.39	0.74	7.23E-04	0.14

**Table 3.8** Validation of protein and cellulose + lignin content retrieval for fresh and dry leaves using stepwise multiple linear regression.

	Protein (g/cm <sup>2</sup> )			Cellulose + lignin (g/cm <sup>2</sup> )		
	$R^2$	RMSE	NRMSE	$R^2$	RMSE	NRMSE
Fresh leaves	0.66	2.02E-04	0.12	0.83	3.91E-04	0.09
Dry leaves	0.74	2.32E-04	0.12	0.79	4.50E-04	0.08



**Figure 3.10** Measured versus estimated protein content (Cp, g/cm<sup>2</sup>) and cellulose + lignin content (Ccl, g/cm<sup>2</sup>) for fresh leaves using PROSPECT-5 inversion with recalibrated specific absorption coefficients and stepwise multiple linear regression (SMLR).



**Figure 3.11** Measured versus estimated protein content (Cp, g/cm<sup>2</sup>) and cellulose + lignin content (Ccl, g/cm<sup>2</sup>) for dry leaves using PROSPECT-5 inversion with recalibrated specific absorption coefficients and stepwise multiple linear regression (SMLR).

### 3.4 Discussion

The leaf protein and cellulose + lignin content were moderately well estimated for both fresh and dry leaves with the recalibrated specific absorption coefficients when PROSPECT-5 was inverted. Comparable accuracies were obtained for estimating cellulose + lignin on fresh and dry leaves through PROSPECT model inversion and stepwise multiple linear regression, irrespective of the water content. Despite of the low amount of leaf protein in fresh leaves, it can still be retrieved via model inversion at promising accuracies. Estimating leaf protein in fresh leaves, using inversion of a radiative transfer model, has been considered

impossible since the 1990s (Jacquemoud et al. 1996), although empirical methods have achieved relatively good accuracies (Kokaly 2001; Kokaly et al. 2009; Kokaly and Clark 1999; Martin et al. 2008). To our knowledge, this is the first study to prove the feasibility of estimating leaf protein and cellulose + lignin in fresh leaves at a moderate to good accuracy using a radiative transfer model.

The specific absorption coefficients of leaf protein and cellulose + lignin determined on fresh, dry and fresh + dry leaves displayed similar magnitudes (Fig. 3.1, 3.2 and 3.3). The coefficients were also quite consistent across species (Fig. 3.4 and 3.5) which can be noted from a repetitive 10-fold cross validation, in each repetition a different data set of species was randomly selected. The findings demonstrated the robustness of the recalibrated specific absorption coefficients, which holds great potential in the development for PROSPECT model given the difficulties in measuring *in vivo* coefficients. Absorption peaks found in the coefficients were in good agreement with known absorption features from literature regardless of the leaf water status (Fig. 3.1, 3.2 and 3.3, Table 3.4 and 3.5). This holds promise for building a general model across a large variety of species and leaf water. Most of known absorption peaks of protein and cellulose + lignin were displayed well in coefficients determined on dry leaves, while fewer absorption features were found in those determined on fresh leaves. The stronger absorption of leaf water on fresh leaves masks some absorption features of protein and cellulose + lignin, such as 1940, 1980 and 2060 nm. And it also led to the off-center wavelength (5-20 nm) of absorption peaks in coefficients from known absorption features in literature, such as 1510, 1730, 2180 nm for protein, and 1780 nm for cellulose + lignin (Table 3.4 and 3.5).

The recalibrated specific absorption coefficients showed similar features for protein (in dry leaves) and cellulose + lignin (in both fresh, dry and fresh + dry leaves) to a previous study (Jacquemoud et al. 1996). However, our results showed great differences in the behavior of the coefficient of protein for fresh leaves to those reported by Jacquemoud et al. (1996). The difference lays not just in the shape of the coefficient curve but also in its magnitude. Our absorption peaks of protein reached approximately 80 cm<sup>2</sup>/g rather than their value of 20 cm<sup>2</sup>/g. This new insight for protein in fresh leaves provides a solid foundation for estimating leaf protein via model inversion. Such improvements can be largely attributed to the new calibration algorithm proposed by Feret et al. (2008). The improvements of the coefficients were further confirmed when applying them in PROSPECT model for forward simulating leaf reflectance and transmittance. The leaf optical properties for fresh leaves were more accurately modelled, especially

over the ranges 1400 - 1800 nm and 2100 - 2500 nm. This means the constituents of protein and cellulose + lignin could explain the variations observed in leaf optical properties, although they only contribute a relatively small proportion to the dry matter. For dry leaves, a small improvement was noted when modeling leaf reflectance, however, the performance deteriorated when transmittance was modeled. This might be due to the limitation of the PROSPECT model in terms of dry leaves, which does not consider inter-correlations between model parameters (Jacquemoud et al. 1996). For example, as a leaf dries, the leaf structure changes with the lower water content.

The selected spectral subsets generally provided equally or more accurate estimation of cellulose + lignin and protein, compared with using full wavelengths of 1100 - 2500 nm. This confirmed that selected spectral subsets often present better results than using full wavelengths (Darvishzadeh et al. 2008; Meroni et al. 2004; Schlerf and Atzberger 2006). However, the spectral subset (Subset 4) defined based on known absorption features of protein did not guarantee more accurate estimation for protein than cellulose + lignin, and the same phenomenon was observed for a specific spectral subset (Subset 5) for cellulose + lignin. This can be explained by the overlapping absorption features due to their common C-H and O-H bonds (Curran 1989; Fourty et al. 1996), and their overlapping absorption features with other secondary compounds (e.g. hemicellulose, starch) which may be in high content for some plant species. The subset 3 of 2100 - 2300 nm proved to be the most effective spectral range for estimating cellulose + lignin and protein on both fresh and dry leaves. This can be attributed to the less influence caused by water absorption, as well as the high values of specific absorption coefficients in this range.

The performances of PROSPECT model inversion and stepwise multiple linear regression were comparable in retrieving cellulose + lignin for both fresh leaves and dry leaves. However, less accurate estimations were obtained for protein when using model inversion other than SMLR. A reason could be the greater abundance of cellulose + lignin than protein in leaves. The relationship between estimated and measured values from model inversion appeared to be less linear than that from SMLR, which indicated larger uncertainties when retrieving parameters using PROSPECT. The uncertainties can be partly explained by the inherent differences lying in radiative transfer models and statistical models. All parameters were simultaneously retrieved through model inversion, thus the parameters making greater contribution to leaf optical properties, such as leaf structure parameter and EWT, inevitably affects the parameters having less

influence, namely cellulose + lignin and protein. In contrast, the statistical model was developed exclusively for one parameter, which is less sensitive to other parameters.

We made efforts to alleviate the “ill-posed” inverse problem and uncertainties in estimation by selecting specific spectral subsets for cellulose + lignin and protein, but the improvement for retrieving protein was limited. Li and Wang (2011) proposed an approach of retrieving parameters in the order of their contributions to leaf reflectance, which is leaf structure parameter, leaf chlorophyll content, equivalent water thickness and leaf mass per area. In that study, one parameter was retrieved at each step by assigning a specific merit function using its own sensitive wavelengths. They first retrieved leaf structure parameter, then estimated the next parameter using the computed values from previous steps as fixed, and continued until all parameters were obtained. The “step-by-step” inversion approach improved the accuracies of estimated leaf chlorophyll content, equivalent water thickness and leaf mass per area, compared to simultaneously determining all parameters using a single merit function (Li and Wang 2011; Zhao et al. 2014). However, the approach did not work well for estimating cellulose + lignin and protein in this study. Different spectral transformation methods have been used to reduce the influence of water on estimating leaf biochemical constituents in statistical studies, such as first or second derivatives, continuum removal (Curran et al. 2001; Huang et al. 2004; Kokaly and Clark 1999; Mutanga and Skidmore 2004; Mutanga et al. 2005), and water removal through spectral matching (Gao and Goetz 1994; Ramoelo et al. 2011; Schlerf et al. 2010). It is worthwhile to investigate such techniques specifically for inversion of a radiative transfer model.

From a practical perspective, the estimation of leaf biochemical content for fresh leaves could have wide application when using airborne or satellite platforms for sensing green canopy foliage (Jacquemoud et al. 1996). So far, there has been a long-held perception, based on the results of Jacquemoud et al. (1996), that leaf protein cannot be retrieved from fresh green leaf spectra when using radiative transfer models. Our results validate one radiative transfer model, i.e. PROSPECT, for estimating leaf biochemical content for fresh leaves, which confirms the statistical models (Asner et al. 2011; Martin et al. 2008; Mutanga and Skidmore 2004; Skidmore et al. 2010) that have consistently demonstrated leaf biochemical content being successfully retrieved from hyperspectral imagery. The ability to model leaf protein, and *ipso facto* leaf nitrogen content, holds the promise that eventually global maps of foliar nitrogen content could be generated



when hyperspectral satellites such as Sentinel (Malenovsky et al. 2012), EnMAP (Kaufmann et al. 2008) and HypSPIRI (<http://hyspiri.jpl.nasa.gov/>) become operational. In turn, such monitoring would support the development and demonstration of Essential Biodiversity Variables (EBVS), which are used for assessing biodiversity (Pereira et al. 2013).

### 3.5 Conclusions

This study demonstrates, for the first time, that leaf protein and cellulose + lignin can be successfully estimated in fresh leaves at a moderate to good accuracy through inversion of a radiative transfer model. The specific absorption coefficients developed for PROSPECT-5 were recalibrated specifically for protein and cellulose + lignin using the new calibration algorithm proposed by Feret et al. (2008). The recalibrated coefficients were robustly estimated across species and showed a greater accuracy when reconstructing leaf reflectance and transmittance. Absorption peaks in coefficients correspond well with known protein and cellulose + lignin absorption features regardless of the leaf water status. The accuracy of retrieved protein and cellulose + lignin from PROSPECT model inversion was comparable with that from stepwise multiple linear regression. The selected spectral subsets improved the accuracy of estimating leaf protein and cellulose + lignin compared with using full wavelengths.

The use of a radiative transfer model offers a robust approach that should be transferable across species and test sites when compared with empirical methods (Darvishzadeh et al. 2008). The LOPEX dataset tested in this study contained plant species with a large variety of leaf structure, leaf biochemistry and leaf optical properties, further confirming the robustness and transferability of this approach. The statistical models can be used to support the development of radiative transfer models in selecting spectral subsets, which provided robust estimations in model inversion. This study was conducted at leaf level, which could be applied at the canopy level for regional or global mapping if coupled with a canopy reflectance model and suitable hyperspectral satellite imagery. However, the confounding factors, such as canopy structure, solar and viewing geometry, and soil background need to be fully considered before upscaling. Canopy level protein (equivalent to nitrogen) (Jacquemoud et al. 1996) and cellulose + lignin (approximation of carbon) would not only yield valuable information about photosynthesis and net production, but would also allow the C:N ratio to be estimated, which can be used for evaluating the nutrient cycle.



## **Chapter 4**

### **Canopy foliar nitrogen retrieved from airborne hyperspectral imagery by correcting for canopy structure effects<sup>\*</sup>**

---

<sup>\*</sup> This chapter is based on: Wang, Z., Skidmore, A.K., Wang, T., Darvishzadeh, R., Heiden, U., Heurich, M., Latifi, H., and Hearne J. (2016). Canopy foliar nitrogen retrieved from airborne hyperspectral imagery by correcting for canopy structure effects. *International Journal of Applied Earth Observation and Geoinformation*, 54, 84-94.

## **Abstract**

A statistical relationship between canopy mass-based foliar nitrogen concentration (%N) and canopy bidirectional reflectance factor (BRF) has been repeatedly demonstrated. However, the interaction between leaf properties and canopy structure confounds the estimation of foliar nitrogen. The canopy scattering coefficient (the ratio of BRF and the directional area scattering factor, DASF) has recently been suggested for use estimating %N as it suppresses the canopy structural effects on BRF. However, estimation of %N using the scattering coefficient has not yet been investigated for longer spectral wavelengths (> 855 nm). We retrieved the canopy scattering coefficient for wavelengths between 400 and 2500 nm from airborne hyperspectral imagery, and then applied a continuous wavelet analysis (CWA) to the scattering coefficient in order to estimate %N. Predictions of %N were also made using partial least squares regression (PLSR). We found that %N can be accurately retrieved using CWA ( $R^2 = 0.65$ , RMSE = 0.33) when four wavelet features are combined, with CWA yielding a more accurate estimation than PLSR ( $R^2 = 0.47$ , RMSE = 0.41). We also found that the wavelet features most sensitive to %N variation in the visible region relate to chlorophyll absorption, while wavelet features in the shortwave infrared regions relate to protein and dry matter absorption. Our results confirm that %N can be retrieved using the scattering coefficient after correcting for canopy structural effect. With the aid of high-fidelity airborne or upcoming space-borne hyperspectral imagery, large-scale foliar nitrogen maps can be generated to improve the modeling of ecosystem processes as well as ecosystem-climate feedbacks.

## 4.1 Introduction

Foliar nitrogen (N) is a primary regulator of physiological processes, such as photosynthesis, leaf respiration, and transpiration (Evans 1989; Field and Mooney 1986; Reich et al. 1998; Reich et al. 2006). Nitrogen is regarded as a limiting nutrient for plant growth (Heimann and Reichstein 2008; LeBauer and Treseder 2008), and is one of the key plant traits driving stand-level forest productivity (Reich 2012). Nitrogen availability constrains carbon assimilation and, thereby, plays an important role in terrestrial ecosystem carbon dynamics, acting as potential climate feedback (Heimann and Reichstein 2008; Lamarque et al. 2005a; Ollinger et al. 2008; Sievering et al. 2000). As a consequence, foliar nitrogen content has recently been proposed as one of the key essential biodiversity variables (EBVs) for satellite monitoring of progress towards the Aichi Biodiversity Targets (Pereira et al. 2013; Skidmore et al. 2015).

Nitrogen has been retrieved with good accuracy using leaf- and canopy-level hyperspectral data despite the fact that it is only a relatively small constituent (0.2-6.4%) in leaves (Kokaly et al. 2009; Wright et al. 2004). Hyperspectral data provides a time-efficient and cost-effective solution for estimating foliar nitrogen compared to the traditional destructive sampling methods. Previous studies on nitrogen estimates used spectra of ground leaf powder, dry leaves and fresh leaves, also cause variation at leaf and canopy level (Curran 1989; Kokaly and Clark 1999; Martin et al. 2008). Nitrogen has been quantified in forest, grassland, and crops (Inoue et al. 2012; Singh et al. 2015; Skidmore et al. 2010). Empirical techniques are dominant in nitrogen retrieval using hyperspectral data, ranging from vegetation indices (Miphokasap et al. 2012), traditional regression techniques such as stepwise multiple linear regression (Kokaly and Clark 1999) and partial least square regression (Martin et al. 2008), to a number of artificial intelligence methods such as support vector regression, neural network and Bayesian model averaging (Axelsson et al. 2013; Skidmore et al. 2010; Zhao et al. 2013).

Mapping forest canopy foliar nitrogen concentration (%N) at a local level is feasible with the aid of imaging spectrometry. The current lack of higher spatial resolution satellite hyperspectral sensors impedes the mapping of nitrogen at larger scales (e.g. continental or global scale), though the launch of the EnMAP mission planned for 2018 may provide an opportunity. Ollinger et al. (2008) provided a continental-scale map of foliar nitrogen by utilizing a statistical relationship between %N and the canopy bidirectional reflectance factor (BRF)

for near-infrared (NIR) wavelengths (800 - 850 nm) for temperate and boreal forests. The findings were encouraging for the remote sensing community, because they allow the estimation of leaf nitrogen across larger extents and provide the option for frequent updates using data from broadband satellite data such as MODIS (Ollinger et al. 2008).

Following the work of Ollinger et al. (2008), Knyazikhin et al. (2013a) applied the physically-based spectral invariants theory (Huang et al. 2007; Knyazikhin et al. 2011) to interpret the process of radiation transfer from leaves and canopies, and explicitly analyzed the coupling between canopy structure and leaf biochemistry in driving variations in canopy BRF. Knyazikhin et al. (2013a) claimed that the significant positive relationship between %N and BRF in the NIR domain (Ollinger et al. 2008) should in fact be attributed to the correlation between canopy structure and BRF. Knyazikhin et al. (2013a) further demonstrated that the BRF between 710 and 790 nm is critical for deriving a canopy structure parameter - the directional area scattering factor (DASF). The DASF is not considered to be a specific canopy structure parameter per se such as leaf area index (LAI), stem density, tree height etc., but rather DASF is a parameter that governs how the scattered radiation from a leaf is further transformed through multiple scattering processes, and can be interpreted as canopy BRF if it is assumed that foliage does not absorb radiation (Knyazikhin et al. 2013a). This was the first study to illustrate the physical interaction between leaf albedo and canopy BRF, which had been neglected in previous studies on hyperspectral remote sensing of leaf biochemical constituents (Knyazikhin et al. 2013a; Ustin 2013).

Although estimation of canopy foliar %N has been tested by using the ratio of BRF and DASF spectra (canopy scattering coefficient,  $W_\lambda$ ) after suppressing the impact of canopy structure (Knyazikhin et al. 2013a), the analyses were restricted to using information from each wavelength between 423 and 855 nm (Knyazikhin et al. 2013a; Knyazikhin et al. 2013b). Canopy foliar %N estimates using the visible spectral region rely on the well-known correlation between chlorophyll and nitrogen (Evans 1989; Field and Mooney 1986; Kokaly et al. 2009). However, the capability of estimating %N using the scattering coefficient based on longer wavelengths from 855 to 2500 nm was not addressed, despite the fact that many nitrogen absorption bands are located in the shortwave infrared regions (SWIR, > 1100 nm) (Curran 1989; Fourty et al. 1996; Kokaly et al. 2009).

Continuous wavelet analysis has emerged as an effective tool in remote sensing – being applied for image fusion, image segmentation, as well as quantifying leaf-level biochemical parameters such as leaf mass per area (Cheng et al. 2014a), leaf chlorophyll (Blackburn and Ferwerda 2008), leaf water content (Cheng et al. 2011; Ullah et al. 2012), and leaf nitrogen (Ferwerda and Jones 2006). The technique has rarely been tested using an imaging spectrometer except in a recent study on detecting water content (Cheng et al. 2014b). Therefore, the feasibility of using continuous wavelet analysis in quantifying canopy foliar %N from airborne hyperspectral data is an interesting challenge.

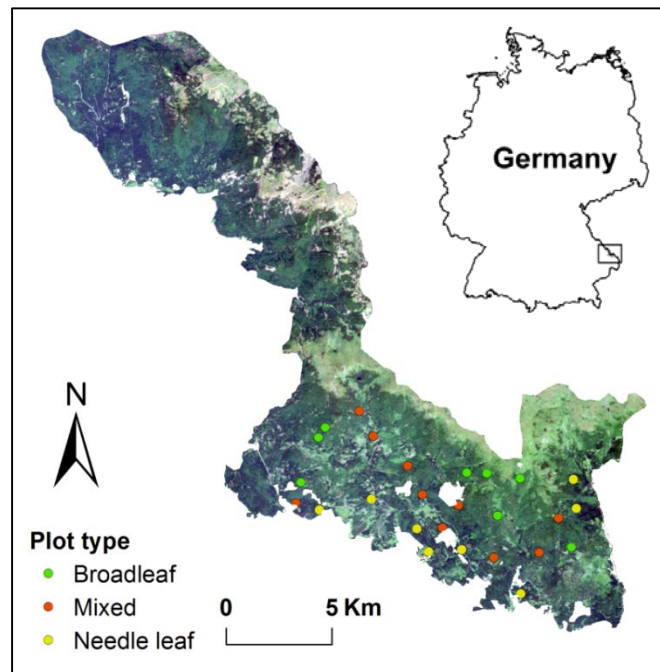
Here, we investigated to what extent canopy foliar %N may be retrieved by disengaging the canopy structural impact from the canopy bidirectional reflectance factor (BRF). We derived the directional area scattering factor (DASF) based on spectral invariant theory, and calculated the canopy scattering coefficient to correct for canopy structural effects. A continuous wavelet transformation was performed on the canopy scattering coefficient. We aimed to (1) test if the scattering coefficient derived from airborne hyperspectral data in the spectral range from 400 to 2500 nm retains information required to estimate %N once canopy structure effects are suppressed; (2) evaluate the performance of continuous wavelet analysis approach in %N prediction and compare the wavelet approach to the widely used partial least squares regression approach; (3) identify the spectral regions most sensitive to %N variations in the canopy scattering coefficient.

## 4.2 Materials and Methods

### 4.2.1 Field data

The study area is located in the southern region of the Bavarian Forest National Park (BFNP) (49° 3' 19" N, 13° 12' 9" E) in southeastern Germany (Fig. 4.1). The BFNP has a total area of 24,218 hectares. The bedrock is primarily composed of gneiss and granite. Soils weathered from these parent materials are naturally acidic and low in nutrients. The main soil types are brown soils, loose brown soils and podzol brown soils. Elevation ranges from 600 to 1453 m. The Bavarian Forest lies in the temperate zone and the climate is subject to both maritime and continental influences. Total annual precipitation is between 1200 and 1800 mm and the mean annual temperature is 5.1°C in the valleys, 5.8°C on hillsides and 3.8°C in the higher montane zones (Heurich et al. 2010). Dominant tree species are Norway spruce (*Picea abies*) (67 %) and European beech (*Fagus sylvatica*)

(24.5 %), with some white fir (*Abies alba*) (2.6 %), sycamore maples (*Acer pseudoplatanus*) (1.2 %) and mountain ash (*Sorbus aucuparia*) (3.1 %) (Heurich et al. 2010).



**Figure 4.1** Location of the study area in the Bavarian Forest National Park, Germany and the sampling plots.

Fieldwork was conducted from mid-July to mid-August, 2013, using a stratified random sampling strategy. The study region was stratified into pure broadleaf, pure needle leaf, and mixed forest areas based on a recent land cover map provided by the Department of Conservation and Research of the BFNP. The mixed forest includes both broadleaf and needle leaf species and their fractions vary across the plots. In order to incorporate variation in canopy structure, we sampled 8 broadleaf, 8 needle leaf and 10 mixed forest plots, which resulted in a total of 26 plots. Each plot was 30 m × 30 m in size, and a Leica GPS 1200 (Leica Geosystems AG, Heerbrugg, Switzerland) was used to record the center location of each plot with an RMSE of approximately 1 m. Within each plot, depending on its homogeneity, one to three trees of each dominant overstory species were selected for sampling. Each sample was composed of at least 20 leaves taken from the branches of an individual tree. The branches of sunlit leaves were shot down from the top of each selected tree using a crossbow. Leaf samples were



stored in zip-lock plastic bags with wet paper towels, and placed in a cooler with ice before transportation to the laboratory for further measurement within 4 hours (Foley et al. 2006). The leaf mass per area (LMA,  $\text{g}/\text{cm}^2$ ), leaf water content ( $\text{g}/\text{cm}^2$ ) and leaf nitrogen concentration (% leaf dry mass) were measured in the laboratory. Further details on the analysis can be found in Wang et al. (2015b).

Within each plot, the canopy structural parameters such as tree height, diameter at breast height (DBH), and number of trees of each dominant species were recorded to calculate species fraction of canopy leaf biomass based on published allometric equations for each species (Gower et al. 1993; Widlowski et al. 2003). Plot-level leaf nitrogen concentration was calculated as the mean leaf nitrogen concentration for each species, weighted by species canopy mass fraction (Knyazikhin et al. 2013a; Martin et al. 2008). The species canopy mass fraction was calculated using the species fraction of canopy leaf biomass and LMA. The plot-level canopy dry matter content ( $\text{kg}/\text{m}^2$ ) was calculated as product of LAI and the mean LMA for each species, weighted by the species canopy leaf area fraction. The plot-level canopy water content ( $\text{kg}/\text{m}^2$ ) was calculated as product of LAI and the mean leaf water content for each species, weighted by the species canopy leaf area fraction. The species canopy leaf area fraction was calculated using the species fraction of canopy leaf biomass and the specific leaf area (the inverse ratio of LMA).

The leaf area index (LAI) for each plot was calculated from five upward-pointing hemispherical photographs collected from the plot center as well as from 10 m from the center in each diagonal direction. The images were acquired using a Canon 5D equipped with a fisheye lens leveled on a tripod at around breast height (1.3 m above the ground) near dawn or dusk. Two-corner classification was applied on the obtained images, and combined Lang and Xiang clumping correction was used to estimate the LAI as outlined in Woodgate et al. (2015), Macfarlane (2011), and Leblanc et al. (2005).

#### **4.2.2 Airborne hyperspectral data acquisition and processing**

A HySpex campaign to acquire airborne hyperspectral imagery was conducted on 22 July 2013 by the German Aerospace Center (DLR). The HySpex sensor-system consists of two imaging spectrometers with spectral ranges of 400 - 1000 nm (VNIR) and 1000 - 2500 nm (SWIR) and a total of 416 spectral channels, which record the solar radiance reflected at the Earth's surface. The HySpex sensor comprises 160 spectral channels in the VNIR and 256 in the SWIR, with a spectral resolution of 3.7 and 6 nm, and a spatial resolution of 1.65 m and 3.3

m for the VNIR and SWIR, respectively. The HySpex data were recorded between 9:00 and 11:00 at an average flying height of 3000 m above ground level. The data were collected in 19 image strips for the study area with approximately 30% overlap, with each strip covering about 1.1 km  $\times$  11 km. The flight line was run in an almost N–S direction. Most of the image strips were acquired in cloud free conditions and none of the sample plots located in the strips had clouds present.

The image strips were preprocessed by DLR. Image data were converted from DN (digital number) to at-sensor radiance using laboratory radiometric calibration information. Ortho-rectification was performed based on the parametric model/table using attitude and flight path data recorded by a differential global positioning system GPS (DGPS) on-board the aircraft in combination with a digital elevation model (DEM). Atmospheric correction was performed to calibrate image data to surface reflectance (canopy BRF as defined before) using the ATCOR4 model based on atmospheric lookup tables generated from the radiative transfer model MODTRAN4. More information about the ATCOR4 model can be found at <http://www.rese.ch/products/atcor/atcor4/>. The Iterative Adaptive Spatial filtering tool (IAS) followed by the Savitzky Golay filter were applied to the atmospherically corrected reflectance to remove random and systematic noise, respectively (Rogge and Rivard 2010; Schl pfer and Richter 2011).

The BRF values for each field plot were extracted from the corrected strips. A 17 by 17 pixel window (i.e. 27.2 m by 27.2 m) around the center of each plot was used to collect the spectra in the VNIR region for each sample, while a 9 by 9 pixel window (i.e. 28.8 m by 28.8 m) was selected for the SWIR domain. For each window, the mean BRF spectrum was calculated to characterize the reflectance of each plot. In total, the spectra of the 26 sampling plots were extracted over the 400 - 2500 nm. The spectral regions located between 1400-1500 and 1800-1950 nm were disregarded because of strong atmospheric water absorption.

#### **4.2.3 Retrieval of DASF and canopy scattering coefficient from BRF spectrum**

The theory of canopy spectral invariants provides a simple and physically-based parameterization of reflected spectra with the green leaf single scattering spectral albedo ( $\omega_s$ ) and two spectrally invariant parameters, i.e., the recollision ( $p$ ) and the escape ( $R$ ) probability (Huang et al. 2007). Here  $p$  is the probability that a

photon scattered from a phytoelement in a canopy will interact within the canopy again (Smolander and Stenberg 2005).  $R$  is the probability that a scattered photon will escape the canopy in a given direction, which is also known as the directional escape factor (Huang et al. 2007; Knyazikhin et al. 2011).

If the impact of canopy background on canopy reflectance is negligible, the spectral bidirectional reflectance factor (BRF) can be approximated as (Huang et al. 2007; Knyazikhin et al. 2011)

$$BRF_{\lambda}(\Omega) = \frac{\rho(\Omega)i_0}{1-\omega_{\lambda}p} \omega_{\lambda} \quad (4.1)$$

where  $\omega_{\lambda}$  is the leaf albedo,  $i_0$  is the canopy interceptance defined as the portion of incoming photons that collide with a phytoelement for the first time.  $p$  is the photon recollision probability,  $\Omega$  is the view direction, and  $\rho(\Omega)$  describes the directional distribution of the total escape probability  $(1-p)$ . The  $\rho(\Omega)i_0$  is referred to as the aforementioned escape factor  $R$ .

As seen from Eq. (4.1), if the leaf albedo at two or more weakly absorbing wavelengths is known, the  $p$  and  $\rho(\Omega)i_0$  can be retrieved. However, the leaf albedo is not available when interpreting air- or space-borne data. Instead, a transformed leaf albedo can be used as a reference (Latorre-Carmona et al. 2014; Schull et al. 2011).

The radiation scattered by a leaf includes two components, surface reflected radiation and the fraction scattered by the leaf interior. In the spectral range of 710 – 790 nm, the fraction of surface reflected radiation is negligible and the diffuse leaf albedo ( $\tilde{\omega}_{\lambda}$ ) dominates the leaf albedo (Knyazikhin et al. 2013a),

$$\omega_{\lambda} \approx i_L \tilde{\omega}_{\lambda} \quad (4.2)$$

where  $i_L$  is the leaf interceptance defined as the fraction of radiation incident on the leaf that enters the leaf interior, and  $\tilde{\omega}_{\lambda}$  is the transformed albedo defined as the fraction of radiation scattered by the leaf interior given that it interacts with internal leaf constituents (Lewis and Disney 2007).

Empirical (Latorre-Carmona et al. 2014; Schull et al. 2011) and theoretical analyses (Lewis and Disney 2007) suggest that, in the spectral range of 710 – 790 nm, the transformed albedo of a given leaf sample is related to a fixed spectrum, ( $\tilde{\omega}_{0\lambda}$ ), via the spectral invariant relationship as,

$$\tilde{\omega}_\lambda = \frac{1-p_L}{1-p_L\tilde{\omega}_{0\lambda}} \tilde{\omega}_{0\lambda} \quad (4.3)$$

where the wavelength-independent within-leaf recollision probability,  $p_L$ , varies per sample.

Considering the spectral range of 710 – 790 nm, Eq. (4.1), (4.2) and (4.3) lead to (Knyazikhin et al. 2013a)

$$\frac{BRF_\lambda(\Omega)}{\tilde{\omega}_{0\lambda}} = p_1 BRF_\lambda(\Omega) + i_L \rho(\Omega)(1 - p_L)i_0 \quad (4.4)$$

where  $p_1 = p_L + i_L p(1 - p_L)$ .

The ratio between the intercept  $i_L \rho(\Omega)(1 - p_L)i_0$  and  $1 - p_1$  is referred to as the directional area scattering factor (DASF), which contains critical information about canopy structure (Knyazikhin et al. 2013a),

$$DASF = \rho(\Omega) \frac{i_L i_0}{1 - p i_L} \quad (4.5)$$

Note that from Eq. (4.2) follows that DASF coincides with BRF (Eq. 4.1) if the transformed leaf albedo is equal to unity. In other words, DASF would be canopy BRF if foliage was a non-absorbing bi-Lambertian surface.

The parameter, DASF, can be simply retrieved from the BRF spectrum of 710 – 790 nm for vegetation canopies with a non-reflecting background without knowing the leaf albedo (Knyazikhin et al. 2013a). The retrieval of DASF are can be described in four steps: 1) calculate the reference leaf albedo,  $\tilde{\omega}_{0\lambda}$ , in the spectrum 710 - 790 nm by using the PROSPECT model with input parameters (leaf structure parameter: 1.5, chlorophyll content: 16  $\mu\text{g}/\text{cm}^2$ , equivalent water thickness: 0.005  $\text{cm}^{-1}$ , and leaf mass per area: 0.002  $\text{g}/\text{cm}^2$ ); 2) for a given  $BRF_\lambda(\Omega)$  spectrum in the 710 - 790 nm, plot the values of the ratio  $BRF_\lambda(\Omega)/\tilde{\omega}_{0\lambda}$ , vs. the values of  $BRF_\lambda(\Omega)$ , and find a linear relationship; 3) find the slope,  $k$ , and intercept  $b$  from the relationship; and 4) the ratio  $b/(1-k)$  forms the estimate of the DASF parameter.

The ratio of  $BRF_\lambda(\Omega)$  to DASF is an estimate of canopy scattering coefficient  $W_\lambda$  (Knyazikhin et al. 2013a), i.e., the fraction of intercepted radiation that has been reflected from, or diffusely transmitted through the vegetation (Lewis and Disney 2007; Smolander and Stenberg 2005). The scattering coefficient can also be interpreted as the ratio of radiation scattered by leaves and exiting the canopy through gaps to radiation scattered by non-absorbing bi-Lambertian leaves and

exiting the canopy through the same gaps. The scattering coefficient relies on leaf interior and surface properties, and provides an estimate of mean leaf albedo.

With the retrieved DASF, the canopy scattering coefficient was calculated as the ratio of canopy BRF and DASF, and used for estimating canopy foliar %N.

#### 4.2.4 Continuous wavelet analysis

Wavelet analysis is an effective signal processing tool to decompose the original data into different scales (Mallat 1989). The technique has been widely used for classification (Meher et al. 2007; Zhu and Yang 1998), image fusion (Otazu et al. 2005; Shi et al. 2005), and dimensionality reduction of hyperspectral data (Bruce et al. 2002; Kaewpijit et al. 2003). Recently, studies have estimated vegetation biophysical and biochemical parameters using hyperspectral data, such as leaf chlorophyll, water content, leaf mass per area and nitrogen (Blackburn and Ferwerda 2008; Cheng et al. 2011; Cheng et al. 2014a; Ferwerda and Jones 2006).

Wavelet transformation can be divided into discrete wavelet transform (DWT) and continuous wavelet transform (CWT) (Bruce et al. 2001). DWT is often used for data compression and feature reduction, and the transformed coefficients of DWT are difficult to interpret, needing inverse transformation for comparison with the original reflectance (Bruce et al. 2001; Cheng et al. 2011). However, the CWT coefficients are interpretable, and can be directly compared with the original spectra to provide information of the shape and position of absorption features in vegetation (Blackburn and Ferwerda 2008; Cheng et al. 2011). The continuous wavelet transform has been used for quantifying leaf water, chlorophyll and dry matter content with leaf reflectance, but studies focusing on canopy level applications are rare (Cheng et al. 2014b). In this study, CWT was adopted to estimate canopy foliar nitrogen using airborne hyperspectral data.

CWT is a linear transformation, which uses a mother wavelet function to convert hyperspectral reflectance  $f(\lambda)$  ( $\lambda = 1, 2, \dots, n$ ,  $n$  is the number of wavebands) into a set of coefficients. The continuous wavelets  $\psi_{a,b}(\lambda)$  can be formulated by scaling and shifting the mother wavelet  $\psi(\lambda)$  (Bruce et al. 2001):

$$\psi_{a,b}(\lambda) = \frac{1}{\sqrt{a}} \psi\left(\frac{\lambda-b}{a}\right) \quad (4.6)$$

where  $a$  and  $b$  are real numbers. The scaling factor ' $a$ ' defines the width of the wavelet and shifting factor while ' $b$ ' determines the position. The CWT output

can be defined by convolving hyperspectral reflectance and the continuous wavelets as (Bruce et al. 2001):

$$W_f(a, b) = \langle f, \psi_{a,b} \rangle = \int_{-\infty}^{+\infty} f(\lambda) \psi_{a,b}(\lambda) d\lambda \quad (4.7)$$

The CWT coefficients ( $W_f(a_i, b_j)$ ,  $i = 1, 2, \dots, m$ ,  $j=1, 2, \dots, n$ ) construct a two-dimensional scalogram ( $m \times n$  matrix) with one dimension of scale (1, 2, ...,  $m$ ) and the other of wavelength (1, 2, ...,  $n$ ). Each scale component is comparable with the reflectance since both of them have the same length. The low and high scale components respectively capture the characteristics of narrow absorption features and the overall shape of spectra (Blackburn and Ferwerda 2008; Ferwerda and Jones 2006). The coefficients scalogram can be used to quantify the variations in shape and depth of spectra due to different amounts of leaf nitrogen.

The absorption features of leaf parameters may be approximated by a Gaussian function (Miller et al. 1990), therefore the Mexican Hat, known as the second derivative of the Gaussian function, was used as the mother wavelet basis (Torrence and Compo 1998). The Mexican Hat effectively supports a range of  $[-5, 5]$  for the scale  $a=1$  and a range  $[-5a_1, 5a_1]$  for the scale  $a=a_1$  (Du et al. 2006). Dyadic scales ( $2^1, 2^2, 2^3, \dots$ , and  $2^8$ ) were used for decomposing the reflectance, instead of a continuous scale ( $i = 1, 2, \dots, 8$ ) to reduce the dimensions of the scalograms, computation and data volume. For simplicity, the dyadic scales ( $2^1, 2^2, 2^3, \dots$ , and  $2^8$ ) in the scalograms were represented as scales  $[1, 2, 3, \dots$  and  $8]$  in the following sections. As suggested by Cheng et al. (2014b), the first three scales of decompositions were not included in the feature selection because they mainly capture the noise in the airborne hyperspectral data.

The continuous wavelet analysis was performed using the MATLAB software package (The MathWorks, Inc.).

The procedure for selecting wavelet features is described as follows. Firstly, the wavelet power (wavelet scalogram) was calculated for all spectra using continuous wavelet transform. The wavelet power is a function of scale and wavelength. Secondly, a correlation scalogram was generated by calculating the coefficient of determination ( $R^2$ ) for the correlation between canopy foliar %N and the wavelet power at each scale and wavelength. A larger value of  $R^2$  indicates a high sensitivity of wavelet features to canopy foliar %N. Thirdly, the features were sorted in descending order of  $R^2$  and only those with the top 5%

values were kept to define the feature regions in the scalogram. Lastly, an individual feature with the highest  $R^2$  was selected from each spectral region. The selected features were used as predictor variables to estimate %N and expressed in wavelengths using the nm scale. Due to the close correlation between %N and canopy dry matter content ( $r^2 = 0.67$ ), we also selected the most significant wavelet features correlated with dry matter to evaluate the effect of dry matter on canopy reflectance variability.

#### 4.2.5 Partial least squares regression

Partial least squares regression (PLSR) transforms the original data to a set of independent variables, which is useful when the number of independent variables is much larger than that of dependent variables (Geladi and Kowalski 1986). Compared with multiple linear regression models, PLSR avoids the problem of co-linearity of variables which is inherent when using hyperspectral data. PLSR has been widely used in the remote sensing community for predicting parameters such as nitrogen, cellulose, lignin, LMA, and LAI (Asner et al. 2011; Darvishzadeh et al. 2011; Martin et al. 2008). To avoid overfitting, the number of variables selected by PLSR is limited to four in this study. The canopy scattering coefficient was used in the PLSR for retrieving canopy foliar %N.

#### 4.2.6 Calibration and validation of the predictive models

All 26 samples were used for the selection of wavelet features. To test the robustness of the selected wavelet features, a “leave-one-out” cross validation procedure of iteratively choosing 25 samples with one sample left was performed for the selection of wavelet features. The selected wavelet features using 25 samples were consistent with those selected by pooled samples. Simple linear regression models were established for canopy foliar %N using a single selected wavelet feature, and multiple linear regression models were also developed using a combination of selected wavelet features.

Both the wavelet features based models and PLSR models were validated using a “leave-one-out” cross-validation, with the final model developed using all samples. For each dependent variable, a model was developed using independent variables from 25 samples, which was then used to estimate the value of the left-out sample. The procedure was repeated 26 times in order to obtain estimates for all samples. The cross-validated coefficient of determination ( $R^2_{cv}$ ), root mean square error ( $RMSE_{cv}$ ), and normalized RMSE ( $NRMSE_{cv} = RMSE_{cv} / \text{range}$ ) between predicted and measured values were calculated to evaluate the

performance of the regression models (Schlerf et al. 2005). We adopted the leave-one-out cross validation approach since it is a well-established method of validation of predictive models in remote sensing studies (Martin et al. 2008; Ollinger and Smith 2005; Wang et al. 2015b). It offered an efficient way of maximizing the use of a relatively small dataset (Pellissier et al. 2015), and enabled a model to be calibrated and validated using all available data (Elisseeff and Pontil 2002). The cross-validated RMSE is a good indicator of the accuracy of the model in predicting unknown samples, because the predicted samples are different from the samples used to build the model (Schlerf et al. 2005).

## **4.3 Results**

### **4.3.1 Characteristics of canopy foliar %N**

Variation in canopy foliar %N was observed in broadleaf, needle leaf and mixed forest plots (Table 4.1). The mean canopy foliar %N in these three types of plots was 2.77, 1.55 and 1.69, respectively, similar to measurements from the Bavarian Forest National Park agency (unpublished data). The pure broadleaf plots had the largest range in %N, namely from 2.18 to 3.29, and largest coefficient of variance (CV), 0.139. The %N in needle leaf plots showed the smallest range (0.24), and consequently, mixed forest provided an intermediate range and CV for %N. The statistics of canopy structural parameters are listed in Table 4.1. The mean of canopy closure and the mean normalized difference vegetation index (NDVI) in broadleaf, needle leaf and mixed forest plots meet the assumption of using the spectral invariants theory in dense vegetation canopies (Knyazikhin et al. 2013a).



**Table 4.1** Statistics of *in situ* canopy foliar %N and canopy structural parameters (sample size=26)

	Canopy foliar %N	LAI	Stem density (n/ha)	Tree height (m)	Canopy closure (%)	NDVI*
<i>Needle leaf</i>						
Min	1.45	3.61	278	15.3	73	0.84
Max	1.69	5.14	1400	30.7	86	0.86
Mean	1.55	4.21	841	24	80	0.85
<i>Broadleaf</i>						
Min	2.18	2.85	244	12.3	76	0.91
Max	3.29	3.24	1022	21.6	89	0.92
Mean	2.77	3.00	675	17.8	83.3	0.91
<i>Mixed</i>						
Min	1.48	2.79	333	11	70	0.85
Max	2.15	4.55	1711	31.9	89	0.92
Mean	1.69	3.63	754	21.1	81.5	0.88

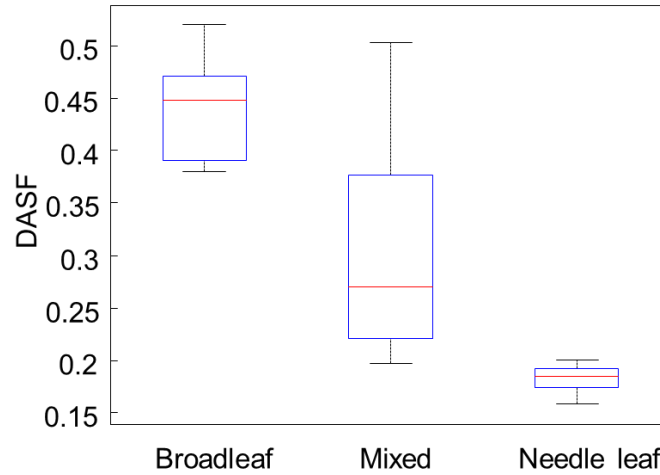
\*NDVI is calculated as the difference between HySpex plot BRFs at 855 nm and 655 nm normalized by their sum.

#### 4.3.2 DASF, canopy BRF spectrum and canopy scattering coefficient

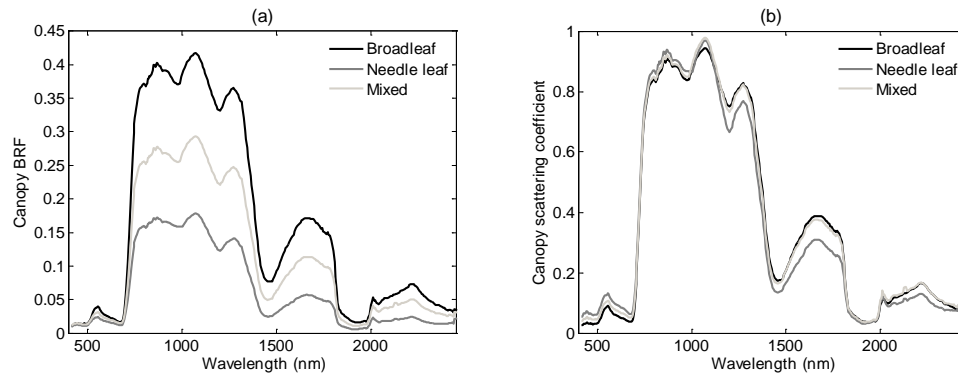
The DASF displayed distinct differences across the broadleaf, needle leaf and mixed forest plots (Fig. 4.2). Broadleaf plots had the largest values of DASF ranging from 0.38 to 0.52, and needle leaf plots produced the lowest values varying from 0.16 to 0.20. Mixed forest yielded intermediate values from 0.20 to 0.50.

As shown in Fig. 4.3 (a), the mean BRF of broadleaf plots had the highest values over the whole spectrum, while that of needle leaf plots had the lowest. The mean BRF of mixed plots gave intermediate values, dependent on the fraction of broadleaf species in those plots. The mean canopy scattering coefficient of the three types of plots showed different behavior across the whole spectral regions (i.e. 400 – 2500 nm) (Fig. 4.3 (b)). After correcting for the canopy structural effects, the scattering coefficient of broadleaf plots was lowest over the spectrum from 400 to 1000 nm. The distinct differences observed with the BRF spectrum among the three plot types over wavelengths between 800 - 1000 nm disappeared.

Furthermore, the scattering coefficient of needle leaf deviated from the other two types of plots in the spectral interval 1200 - 2500 nm. The needle leaf, though having a lower nitrogen concentration, absorbs more photons because of its high dry matter content (Zarco-Tejada et al. 2004), thus resulting in a lower albedo.



**Figure 4.2** Boxplots of the directional area scattering factor (DASF) retrieved based on spectral invariants theory in the broadleaf, needle leaf, and mixed forest plots.



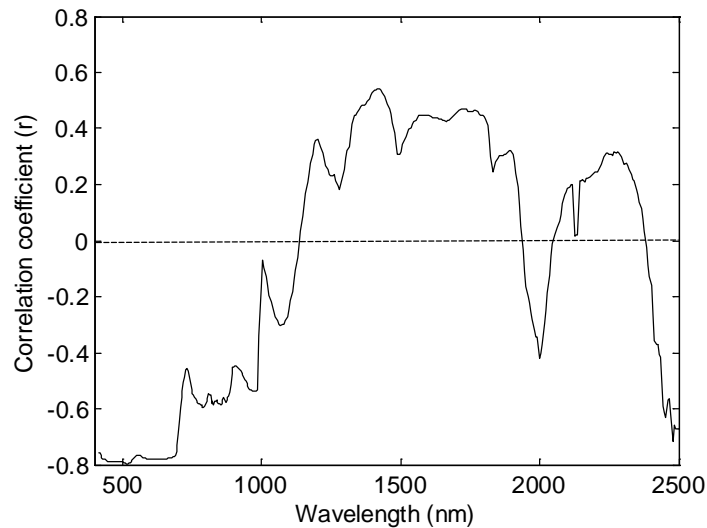
**Figure 4.3** Comparison for (a) mean canopy BRF (left) and (b) mean canopy scattering coefficient (right) of the broadleaf, needle leaf, and mixed forest plots.

#### 4.3.3 Correlation between canopy foliar %N and canopy scattering coefficient

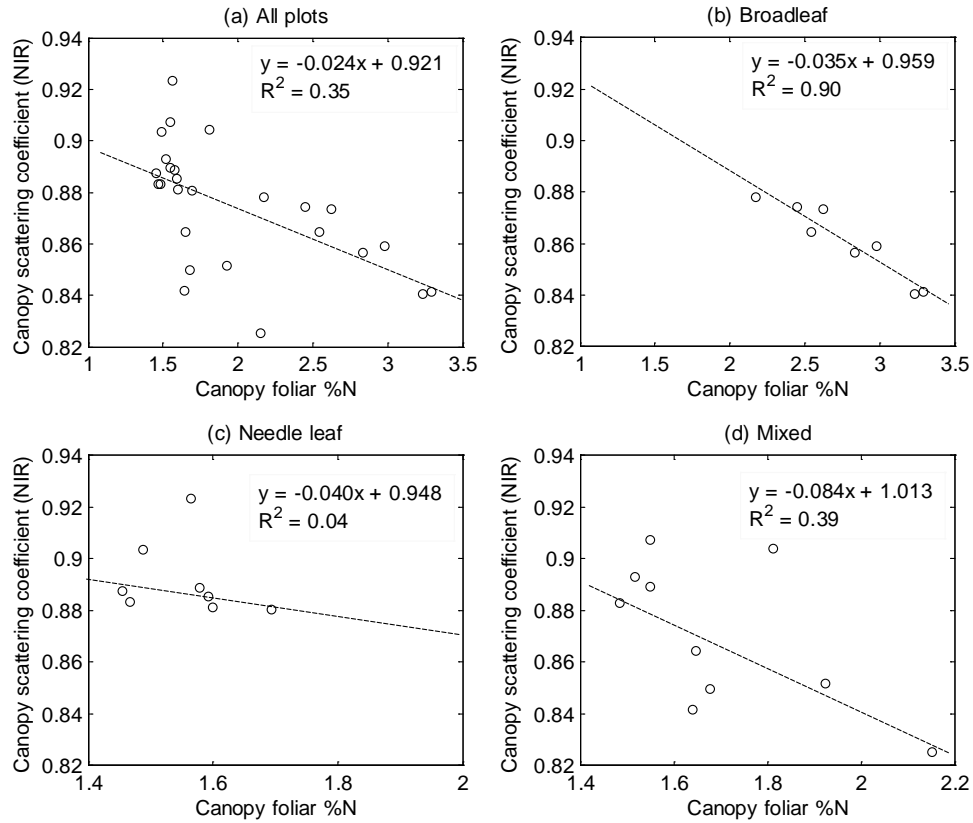
We calculated the correlation between foliar %N and the canopy scattering coefficient over the whole spectrum from 400 to 2500 nm, as shown in Fig. 4.4. Negative correlations were observed in the 400 – 1200 nm interval. In contrast, %N was positively correlated with the canopy scattering coefficient in

the 1200 – 1800 nm and 2100 – 2400 nm spectral intervals (Fig. 4.4). The correlation in these spectral regions ( $r < 0.55$ ) was, however, much weaker than in the 400 - 700 nm interval ( $r < -0.75$ ).

We noted a negative relationship between foliar %N and the canopy scattering coefficient over the NIR spectral range (800 - 850 nm) for the pooled plots ( $R^2 = 0.35$ ,  $p < 0.01$ ) (Fig. 4.5). A stronger correlation was revealed when solely considering broadleaf plots ( $R^2 = 0.90$ ,  $p < 0.001$ ). A weaker relationship was found for mixed forest plots ( $R^2 = 0.39$ , not significant). And the negative relationship between foliar %N and the canopy scattering coefficient with regard to needle leaf plots was not significant ( $R^2 = 0.04$ ).



**Figure 4.4** Correlation coefficient ( $r$ ) between the canopy foliar %N and canopy scattering coefficient over 400 - 2500 nm.



**Figure 4.5** Canopy foliar %N vs. canopy scattering coefficient in the NIR spectral wavelengths (800-850 nm), (a) all plots, (b) broadleaf, (c) needle leaf, and (d) mixed forest plots.

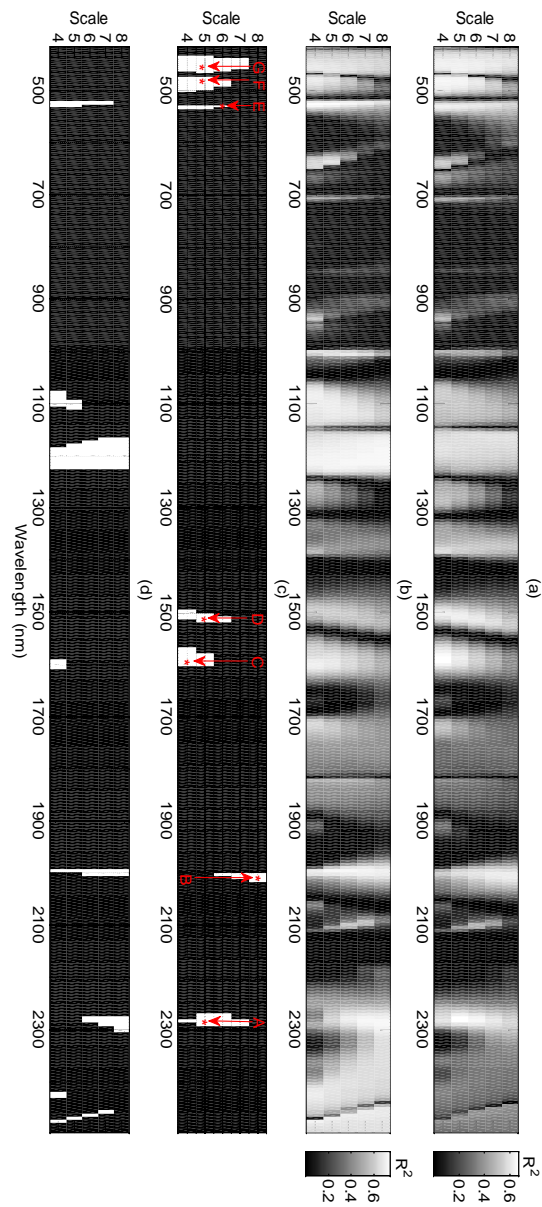
#### 4.3.4 Wavelet features sensitive to %N

The correlation ( $R^2$ ) between wavelet power and canopy foliar %N and canopy dry matter content are displayed in Fig. 4.6(a) and Fig. 4.6(b), respectively. Fig. 4.6(c) and Fig. 4.6 (d) delineate the top 5% wavelet features correlated with %N and canopy dry matter content, respectively. The correlation between wavelet power and canopy water content was similar to that between wavelet power and dry matter (results not shown), due to the high correlation between water and dry matter ( $r^2 = 0.93$ ).

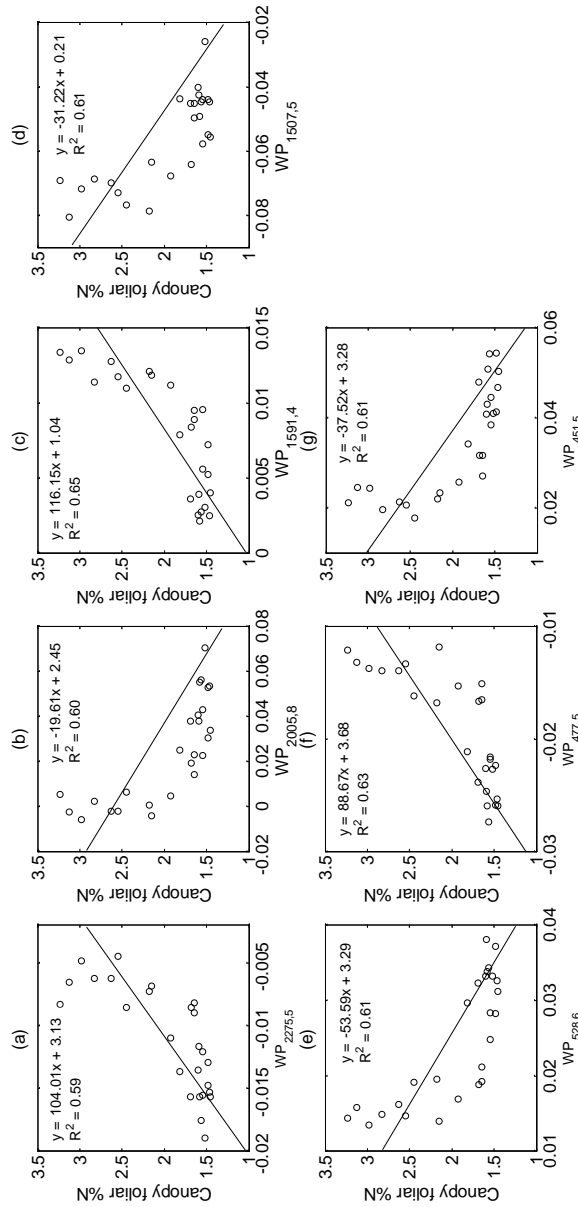
Seven wavelet features most sensitive to %N variations were identified in each of the delineated regions (Fig. 4.6(c) and Table 4.2). Statistically significant relationships were observed between %N and the wavelet features with  $R^2$  ranging from 0.59 to 0.65 ( $P < 0.0001$ ) (Fig. 4.7). The most significant

relationship between a single wavelet feature and %N was obtained by the wavelet feature WP<sub>1591,4</sub>. After testing all combinations of the seven wavelet features, the combination of WP<sub>2005,8</sub>, WP<sub>1507,5</sub>, WP<sub>528,6</sub> and WP<sub>477,5</sub> generated a significant relationship with leaf %N ( $R^2 = 0.75$ ,  $P < 0.0001$ , not shown).

We observed that the four selected wavelet features A-D (Table 4.2) were from the SWIR spectral regions, a region which is largely influenced by dry matter and water absorption. Therefore, we regressed each of these four wavelet features against %N and canopy dry matter content (as the two independent variables) and calculated the standardized beta coefficients of each variable (Bring 1994; Landis 2005) in order to compare the relative contribution of each variable to the variation in the wavelet feature. We also regressed each of these four wavelet features against %N and canopy water content. Nitrogen, canopy dry matter content, and canopy water content were not involved in one multiple linear regression because of the high correlation between dry matter and water content ( $r^2 = 0.93$ ). Results showed that nitrogen contributed more to the variations of the following wavelet features WP<sub>1507,5</sub> and WP<sub>2275,5</sub> than dry matter and water due to having higher standardized beta coefficients (results not shown). Dry matter and water, on the other hand, made a larger contribution to the wavelet features WP<sub>1591,4</sub> and WP<sub>2005,8</sub> than nitrogen (results not shown).



**Figure 4.6** Correlations scalograms representing the correlation ( $R^2$ ) between wavelet features and (a) canopy foliar %N and (b) canopy dry matter content. (c) and (d) delineate the top 5% wavelet features correlated with %N and canopy dry matter content, respectively. The seven selected wavelet features (A-G) sensitive to %N in (c) are detailed in Table 4.2. The wavelet power was derived from the canopy scattering coefficient.

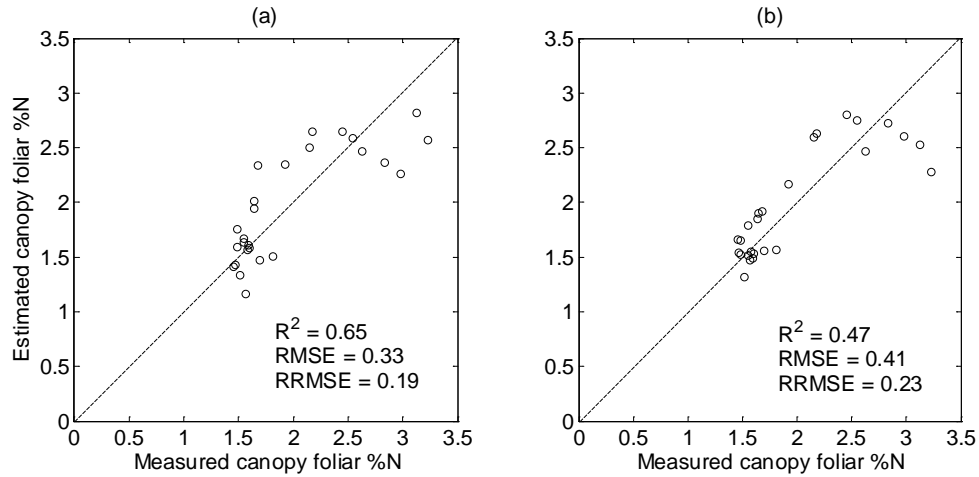


**Figure 4.7** Relationships between canopy foliar %N and wavelet features (a)  $WP_{2275.5}$ , (b)  $WP_{2005.8}$ , (c)  $WP_{1591.4}$ , (d)  $WP_{1507.5}$ , (e)  $WP_{528.6}$ , (f)  $WP_{477.5}$  and (g)  $WP_{451.5}$ , derived from the canopy scattering coefficient. Data in this figure correspond to pooled samples. All relationships are statistically significant ( $P < 0.0001$ ).

#### 4.3.5 Estimated canopy foliar %N using wavelet features and PLSR

The regression model based on wavelet feature  $WP_{1591.4}$  derived from the canopy scattering coefficient provided the most accurate estimation of canopy foliar %N with a cross-validated  $R^2$  ( $R^2_{CV}$ ) of 0.59 and RMSE (RMSE<sub>CV</sub>) of 0.36 (Table 4.2). The regression model using the combination of wavelet features,  $WP_{2005.8}$ ,  $WP_{1507.5}$ ,  $WP_{528.6}$  and  $WP_{477.5}$ , improved the estimation accuracy of leaf %N by

yielding a  $R^2_{CV}$  of 0.65 and a  $RMSE_{CV}$  of 0.33 (Fig. 4.8(a) and Table 4.2). In comparison, the PLSR provided a less accurate estimation of %N ( $R^2_{CV} = 0.47$ ,  $RMSE_{CV} = 0.41$ ) (Fig. 4.8(b)).



**Figure 4.8** Canopy foliar %N estimated from (a) regression models using the combination of wavelet features  $WP_{2005,8}$ ,  $WP_{1507,5}$ ,  $WP_{528,6}$  and  $WP_{477,5}$  derived from the scattering coefficient, and (b) partial least squares regression model.  $R^2$  and RMSE in the figure are cross-validated values.

**Table 4.2** Performance of regression models for estimating canopy foliar %N using different wavelet features from the canopy scattering coefficient. Related features can be referred to Curran (1989).

Code	Wavelet features	$R^2_{CV}$	$RMSE_{CV}$	$NRMSE_{CV}$	Features related to
A	$WP_{2275,5}$	0.52	0.39	0.22	Protein and nitrogen at 2300 nm
B	$WP_{2005,8}$	0.53	0.39	0.22	Water at 1940 nm
C	$WP_{1591,4}$	0.59	0.36	0.20	Starch and sugar at 1580 nm
D	$WP_{1507,5}$	0.54	0.39	0.22	Protein and nitrogen at 1510 nm
E	$WP_{528,6}$	0.54	0.38	0.22	-
F	$WP_{477,5}$	0.56	0.38	0.21	Chlorophyll b at 460 nm
G	$WP_{451,5}$	0.55	0.38	0.21	Chlorophyll b at 460 nm
	$WP_{\text{combo}} (B, D, E, F)$	0.65	0.34	0.19	-

## 4.4 Discussion

Hyperspectral remote sensing of foliar nitrogen may be confounded by canopy structure, causing variation in the canopy BRF (Knyazikhin et al. 2013a). This



study derived the canopy scattering coefficient by correcting for canopy structure effects on the canopy BRF over the full spectrum from 400 to 2500 nm using airborne hyperspectral data. Our results confirmed the feasibility of accurately estimating canopy foliar %N ( $R^2 = 0.65$ , RMSE = 0.33, Table 4.2) from the scattering coefficient using a wavelet based approach. This wavelet based approach yielded a higher accuracy than a standard PLSR approach. The wavelet features most sensitive to nitrogen variations were selected from the visible and shortwave near-infrared spectral regions, which relate to absorption features of chlorophyll, protein, and dry matter.

A spectral invariants variable, DASF, serves as a parameter for characterizing canopy structure. As seen in Fig. 4.2, the different canopy structure of the broadleaf, needle leaf and mixed forest plots were reflected in the DASF values. For a non-absorbing canopy ( $\omega_\lambda \equiv 1$ ), the DASF equals the canopy BRF (Eq.(4.1)) (Knyazikhin et al. 2013a). The BRF of a non-absorbing canopy provides an estimate of the fraction of the leaf area inside the canopy that is visible from the outside of the canopy along a given direction. The tendency in the DASF, which is highest for broadleaf and lowest for needle leaf (Fig. 4.2) corresponds with the extent of foliage clumping, i.e., needle leaves in shoots give rise to multiple photon-needle interactions within a shoot (Ollinger 2011), thus reducing the fraction of the leaf area that can be seen by the sensor. In addition, broadleaf stands in general have spherical to ellipsoidal-shaped crowns and needle-leaf stands have ellipsoidal to conical shapes. The smooth and continuous upper surfaces of broadleaf trees also allow more photons to be detected by a nadir viewing sensor (Knyazikhin et al. 2013a).

A negative correlation between %N and the canopy scattering coefficient between 400 and 1200 nm was observed in Fig. 4.4. This phenomena further confirmed that the positive correlation between %N and NIR reflectance reported by Ollinger et al. (2008) was due to the relationship between canopy structure and NIR reflectance (Knyazikhin et al. 2013a). In the spectral region between 400 – 770 nm, pigments such as chlorophyll dominates the absorption (Feret et al. 2008). Thus, the broadleaf trees with higher nitrogen concentration (more chlorophyll) absorbs more photons, resulting in a lower scattering coefficient (Fig. 4.3(b)) and a negative correlation between %N and the scattering coefficient (Fig. 4.4). In contrast, the scattering coefficient was positively correlated with %N in the spectral intervals 1200 – 1800 nm and 2100 – 2400 nm (Fig. 4.4). In these spectral regions, the absorption by dry matter is much stronger than nitrogen. Needles with low %N absorb more photons because of their higher dry matter

content (Zarco-Tejada et al. 2004), which leads to a lower canopy scattering coefficient and a positive correlation with %N (Fig. 4.4).

The feasibility of using canopy scattering coefficient to estimate canopy foliar %N was evaluated by also considering the wavelengths of 855 to 2500 nm. The regression models developed for the wavelet features provided an accurate estimation of %N ( $R^2 = 0.65$ , RMSE = 0.33, Table 4.2). Four wavelet features sensitive to %N were identified in the SWIR spectral regions (Fig. 4.6 (b) and Table 4.2), which could be related to protein and dry matter absorption. Nitrogen contributed more to the variation in the wavelet features  $WP_{1507,5}$  and  $WP_{2275,5}$  than dry matter and water did (Section 3.4), and the two wavelet features captured the narrow-range absorption by protein and nitrogen at 1510 nm and 2300 nm, respectively, as demonstrated in previous studies (Curran 1989; Fourty et al. 1996; Martin et al. 2008). The other two wavelet features identified,  $WP_{1591,4}$  and  $WP_{2005,8}$  were more influenced by dry matter and water than by nitrogen (Section 3.4). The wavelet feature  $WP_{1591,4}$  captured the narrow-range absorption of starch and sugar (dry matter constituents) at 1580 nm (Curran 1989). The wavelet feature  $WP_{2005,8}$  was mainly controlled by the broad water absorption feature centered at 1940 nm (Curran 1989).

Another three wavelet features sensitive to %N were selected in the visible spectral regions (Fig. 4.6 (b) and Table 4.2). The sensitivity of the wavelet features at 451 nm, 477 nm, and 528 nm, could be explained by the correlation between nitrogen with chlorophyll (Evans 1989; Field and Mooney 1986; le Maire et al. 2008). The wavelet features  $WP_{451,5}$  and  $WP_{477,5}$  were located at the edge of the chlorophyll b absorption features at 460 nm (Curran 1989; Curran et al. 2001). The wavelet feature  $WP_{528,6}$  is not directly linked to chlorophyll absorption features, though has been identified as a waveband sensitive for nitrogen estimation in previous studies (Huang et al. 2004; Mutanga et al. 2004). In addition, the closeness of this band to 531 nm, which is used for calculating the photochemical reflectance index (Gamon et al. 1992), might indicate a link between xanthophyll cycle pigments and nitrogen in the photosynthetic process.

Estimation of nitrogen using hyperspectral remote sensing relies on the fact that nitrogen in leaves causes variation in reflectance. Both direct and indirect effects of nitrogen are exerted on reflectance. Since nitrogen is only a relatively small constituent (0.2-6.4%) of leaves (Wright et al. 2004), the direct effect of nitrogen on reflectance through interacting with photons is minimal (Jacquemoud et al. 1996). The absorption features of nitrogen also overlap with those of dry matter

and water. In this study, the two wavelet features  $WP_{1507,5}$  and  $WP_{2275,5}$  were found to be related to nitrogen the most. The indirect effects of nitrogen on reflectance may be attributed to associations between nitrogen, dry matter, water and chlorophyll. For instance, the other five selected wavelet features were subject to the collective effects of the absorption of these constituents. Both the direct and indirect effects of nitrogen on the reflectance are critical for %N estimation, and need to be further evaluated in future research.

Although it has recently been demonstrated that leaf protein and *ipso facto* nitrogen can be estimated in fresh leaves by inversion of a leaf radiative transfer model PROSPECT (Jacquemoud and Baret 1990; Jacquemoud et al. 1996; Wang et al. 2015a), the feasibility of retrieving foliar %N using coupled leaf and canopy models still needs to be investigated, especially given the more complicated factors such as viewing/solar geometry, canopy structure and soil background which all strongly affect the canopy spectra. In this regard, the strategy of combining empirical and radiative transfer models may provide a novel method to model foliar %N using hyperspectral remote sensing. The spectral invariants variable, DASF, may serve as an efficient mechanism to correct for the impact of canopy structure on canopy BRF. The continuous wavelet analysis is capable of identifying the wavelet features that are most sensitive to nitrogen variations and can be used for %N estimation. The continuous wavelet analysis has been shown to be efficient in predicting leaf biochemical parameters such as leaf chlorophyll, leaf mass per area and leaf water content (Blackburn and Ferwerda 2008; Cheng et al. 2011; Cheng et al. 2014a), as well as canopy water content (Cheng et al. 2014b). To our knowledge, this is the first study to evaluate the feasibility of a wavelet based approach to estimate canopy foliar %N using airborne hyperspectral data.

## 4.5 Conclusions

To estimate canopy foliar nitrogen the canopy BRF was evaluated across all wavelengths from 400 – 2500 nm, correcting for canopy structural impacts using the spectral invariants parameter, DASF. DASF is derived based on physical laws and serves as a practical way of characterizing canopy structure. Canopy foliar nitrogen can be accurately estimated from the canopy scattering coefficient, which suppresses the impact of canopy structure. Continuous wavelet analysis led to a more accurate estimation of canopy foliar nitrogen than partial least square regression. Wavelet features sensitive to nitrogen variations were

identified from visible and short near-infrared spectral regions, which relate to the absorption features of chlorophyll, protein, and dry matter.

The decoupling of canopy structure from the canopy BRF is a common challenge in the estimation of all leaf biochemistry, such as leaf water, dry matter, cellulose, lignin, and sun-induced chlorophyll fluorescence (Damm et al. 2015; Knyazikhin et al. 2013a). Earth observation missions such as the Sentinel series (Malenovsky et al. 2012), EnMAP (Guanter et al. 2015) and HypSIRI and DESDynI LiDAR-RaDAR instrument ((NRC) 2007) provide rich global datasets for estimating spectral and canopy structural properties. It has been suggested that nitrogen, as one of the proposed key essential biodiversity variables (EBVs), should be tracked from space for global monitoring by ecologists and space agencies (Pereira et al. 2013; Skidmore et al. 2015). The combination of the spectral invariants theory and continuous wavelet analysis offers an efficient and robust way of hyperspectral remote sensing of canopy foliar nitrogen regardless of the confounding effects of canopy structure. Regional to global scale mapping of foliar nitrogen can facilitate research on evaluating and monitoring biodiversity. In a broader context, foliar nitrogen as an input to ecosystem process models is essential in understanding the nutrient cycle and climatic feedbacks (Singh et al. 2015; Zaehle et al. 2014).

## Chapter 5

### **Mapping leaf and canopy nitrogen content from airborne hyperspectral imagery using coupled leaf-canopy radiative transfer models\***

---

\* This chapter is based on: Wang, Z., Skidmore, A.K., Darvishzadeh, R., and Wang, T. (Submitted). Mapping leaf and canopy nitrogen content by inversion of radiative transfer models from airborne hyperspectral imagery. *Remote Sensing of Environment*.

## **Abstract**

A revised leaf radiative transfer model PROSPECT-5 incorporated the effects of leaf protein and cellulose + lignin on leaf reflectance and transmittance to estimate leaf protein from fresh leaf spectra (Wang et al. 2015a). Leaf protein can be calculated as 6.25 times of leaf nitrogen (Barton 1987; Jacquemoud et al. 1996), which may provide an approach of estimating nitrogen using radiative transfer models. However, the revised leaf model has not been tested for the estimation of leaf nitrogen at the canopy level. In this study, a canopy reflectance model INFORM, coupled with the revised PROSPECT-5 model, was used to successfully retrieve leaf and canopy nitrogen content from airborne hyperspectral imagery. Ecological criteria were applied to the parameterization of the model to reduce unrealistic combinations of input parameters. Global sensitivity analysis showed that leaf protein played a small but distinct role in driving the variation of canopy reflectance in the INFORM model. More accurate estimation was obtained for canopy nitrogen content ( $R^2 = 0.61$ , RMSE = 2.40, NRMSE = 0.22) than leaf nitrogen content ( $R^2 = 0.38$ , RMSE = 0.000040, NRMSE = 0.20). Moreover, inversion techniques, particularly regularized look-up tables, further improved the estimation accuracies compared to the original tables. Our results indicate that leaf and canopy nitrogen content can be retrieved at the canopy level by inversion of INFORM. Our maps of leaf and canopy nitrogen content are the first to be generated using inversion of coupled leaf-canopy models, and the spatial variation of foliar nitrogen appears to be reasonable and consistent with ecological knowledge.

## 5.1 Introduction

Leaf nitrogen is an important parameter related to photosynthesis and net primary production (Evans 1989; Field and Mooney 1986; Reich 2012), which mainly exists in chlorophylls and proteins in the leaf cells (Elvidge 1990; Evans 1983). Estimation of leaf nitrogen improves our understanding of terrestrial ecosystem carbon dynamics and climate models (Lamarque et al. 2005d; Morford et al. 2011; Ollinger et al. 2008). As one of the essential biodiversity variables, foliar nitrogen can be used to assess biodiversity and ecosystem services (Pereira et al. 2013; Skidmore et al. 2015). Accurate retrieval of leaf nitrogen could therefore enhance ecosystem process models that describe ecosystem functioning, since nitrogen is an important input parameter of these models (Zaehle et al. 2014; Zhang et al. 2013).

Leaf nitrogen has been quantified from leaf and canopy spectra as nitrogen absorption features through detection in the narrow-band of continuous information from hyperspectral remote sensing (Cho 2007b), although it is a small constituent of leaf dry weight (0.2% - 6.4%) (Wright et al. 2004). A number of factors confound the retrieval of leaf nitrogen from leaf and canopy spectra. Leaf water is one of the main obstacles for estimating leaf nitrogen using the fresh leaf spectra, because the strong absorption of water masks the spectral features of nitrogen in the short-wave near infrared (SWIR) spectral region (Fourty and Baret 1998; Kokaly and Clark 1999). Other constituents in leaves, such as cellulose, lignin and starch, also overlap with the absorption features of nitrogen in SWIR regions (Curran 1989), which further increases the difficulty of retrieving nitrogen data.

When estimating leaf nitrogen from the canopy spectra, factors such as canopy structure, illumination/viewing geometry, and the background can further decrease our ability to detect nitrogen (Asner 1998; Knyazikhin et al. 2013a; Yoder and Pettigrew-Crosby 1995; Zarco-Tejada et al. 2001). Different approaches have been applied to improve leaf nitrogen estimation by enhancing the absorption features of nitrogen. In addition, reducing the influence of other factors on canopy reflectance has been successfully developed, including spectral transformation such as using first/second derivatives and log transformation of reflectance (Coops et al. 2003; Yoder and Pettigrew-Crosby 1995), continuum removal (Huang et al. 2004; Kokaly and Clark 1999), water removal (Ramoelo et al. 2011; Schlerf et al. 2010), and wavelet analysis (Ferwerda and Jones 2006). Empirical approaches such as vegetation indices (Serrano et al. 2002; Wang et al.

2016a), stepwise multiple linear regression (Kokaly and Clark 1999), partial least squares regression (Lepine et al. 2016; Martin et al. 2008), neural networks (Skidmore et al. 2010), and support vector regression (Axelsson et al. 2013) have been used to establish relationships between spectral data and nitrogen. However, such relationships have been site-, sensor-, date- or species-specific. In addition, for a variety of reasons, the selected wavelengths that are sensitive to nitrogen in different studies are not always consistent (Curran et al. 2001; Huang et al. 2004; Kokaly and Clark 1999).

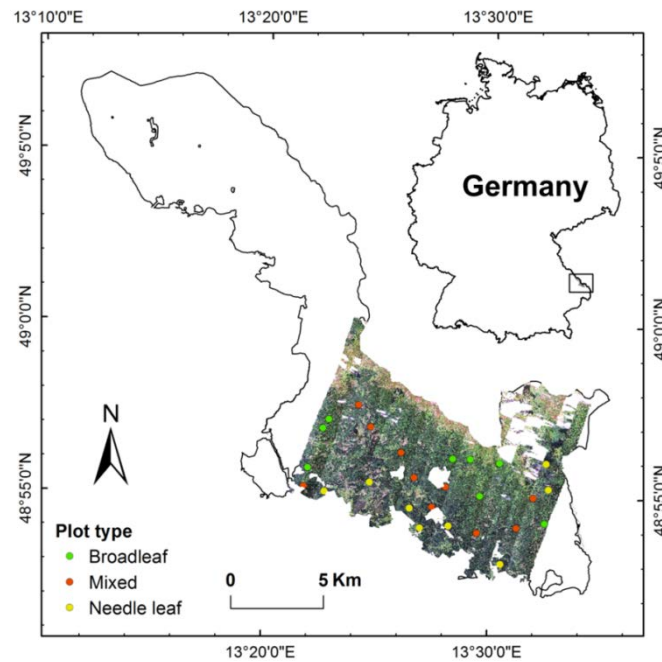
Radiative transfer models (RTMs) offer a conceptual superiority to empirical approaches with respect to transferability and robustness (Darvishzadeh et al. 2011; Jacquemoud and Baret 1990; Schlerf and Atzberger 2006). In these models, the transfer and interactions of electromagnetic radiation inside the canopies are described based on physical laws (Verhoef 1984). Depending on the model type, i.e. leaf or canopy, the absorbing and scattering processes of radiation are incorporated using a range of leaf, canopy and external parameters (Jacquemoud and Baret 1990; Verhoef 1984). It was considered impossible to retrieve leaf nitrogen from fresh leaves using leaf RTMs (Jacquemoud et al. 1996). However, it has recently been shown that leaf protein can be estimated through the leaf optical properties model PROSPECT-5 (Feret et al. 2008), which was revised to incorporate both protein as well as cellulose + lignin compounds as a replacement for the dry matter (Wang et al. 2015a). Leaf protein can be calculated as 6.25 times of leaf nitrogen (Barton 1987; Jacquemoud et al. 1996), which provides the possibility of nitrogen estimation using radiative transfer models. However, this revised leaf model had not yet been assessed for retrieving leaf nitrogen when scaling up to the canopy level by coupling it with a canopy reflectance model.

The canopy reflectance model provides a means of understanding the covariance of leaf and canopy effects in canopy reflectance (Baret et al. 1994; Jacquemoud et al. 2000). There are generally four categories of canopy reflectance models: (1) 1D turbid medium model such as SAILH (Verhoef 1984); (2) geometrical models such as Li–Strahler GO model (Li and Strahler 1985); (3) Monte Carlo ray tracing models such as DART (Gastellu-Etchegorry et al. 1996); and (4) hybrid models such as GeoSail (Huemmrich 2001). To select a proper canopy reflectance model, two factors should be considered (Atzberger 2000; Pinty et al. 2004). The first factor is the realism of simulations with regard to the canopy architecture description, and the second is the invertibility of the model associated with a limited number of input variables. However, both these factors are controversial, and a compromise needs to be sought. The hybrid models benefit from a



combination of 1D turbid medium and geometrical models (GO), which means they are closer to reality and easier to invert. The invertible forest reflectance model INFORM (Atzberger 2000) is an example of a hybrid model that has been successful in retrieving vegetation parameters (Ali et al. 2016a; Schlerf and Atzberger 2006, 2012; Yuan et al. 2015). As a hybrid model, INFORM could therefore offer a compromise between the realism of simulating the canopy and invertibility. We adopted the INFORM model for this study (Wang et al. 2015a).

We aimed to evaluate the performance of scaling-up approaches to estimating leaf and canopy nitrogen content from airborne hyperspectral imagery, by coupling the revised PROSPECT-5 model (Wang et al. 2015a) with a canopy reflectance model, i.e. INFORM. Our specific objectives were: (1) to investigate the feasibility of using a coupled leaf-canopy reflectance model for retrieving leaf and canopy nitrogen content from canopy spectra; and (2) to map the spatial variation of leaf and canopy nitrogen content in a mixed temperate forest.



**Figure 5.1** Location of the study area in the Bavarian Forest National Park, Germany, and the distribution of sample plots with the HySpex image in the background.

## 5.2 Materials and methods

### 5.2.1 Study area and field data

#### 5.2.1.1 Study area

The study area is located in the southern part of the Bavarian Forest National Park (49° 3' 19'' N, 13° 12' 9'' E), Germany (Fig. 5.1). The park has a total area of 24,218 hectares. The geology is dominated by gneiss and granite, and the soils weathered from these parent materials are naturally acid and low in nutrients. The main soil types are brown soils, loose brown soils, and podsol brown soils. The park's elevation ranges from 600 m to 1453 m. The climate is temperate with a total annual precipitation between 1200 mm and 1800 mm and a mean annual temperature of 5.1°C in the valleys, 5.8°C on hillsides, and 3.8°C in the higher montane zones (Heurich et al. 2010). The dominant forest species are Norway spruce (*Picea abies*) (67 %) and European beech (*Fagus sylvatica*) (24.5 %), with some white fir (*Abies alba*) (2.6 %), sycamore maples (*Acer pseudoplatanus*) (1.2 %), and mountain ash (*Sorbus aucuparia*) (3.1 %) (Heurich et al. 2010).

#### 5.2.1.2 Field sampling

Fieldwork was carried out from mid-July to mid-August, 2013, using a stratified random sampling strategy. The study area was stratified into broadleaf, needle leaf, and mixed forest areas based on the land cover data obtained from the Department of Conservation and Research, Bavarian Forest National Park. The mixed forest includes both broadleaf and needle leaf species, but the fractions vary across different areas. We randomly selected 26 plots over the three vegetation types, yielding 8 broadleaf, 8 needle leaf and 10 mixed forest plots. Each plot was 30 m × 30 m in size, and a Leica GPS 1200 (Leica Geosystems AG, Heerbrugg, Switzerland) was used to record the center location of each plot (to an accuracy of approximately 1 m). Within each plot, depending on the species composition, one to three trees of each dominant overstory species were selected for leaf sampling.

Sunlit leaves were collected by shooting small branches from the top canopy of each selected tree with a crossbow. Leaf samples were stored in zip-lock plastic bags with wet paper towels, and placed in a cooler before transportation to the laboratory for further measurement. The leaf mass per area (LMA, g/cm<sup>2</sup>), leaf water content (g/cm<sup>2</sup>, also known as equivalent water thickness (EWT, cm)) and leaf nitrogen concentration (% dry weight) were determined in the laboratory.

Leaf nitrogen content ( $N_{\text{area}}$ , g/cm<sup>2</sup>) was calculated as the product of leaf nitrogen concentration and LMA. More details regarding the sampling and laboratory analysis can be found in Wang et al. (2015b).

### 5.2.1.3 Measurements of canopy structural parameters

A number of canopy structural parameters were collected within each plot, including leaf area index (LAI), stem density, canopy closure, crown diameter and stand height. LAI for each plot was calculated from five upward-pointing hemispherical photographs collected from the plot center as well as 10 m away from the center point in each diagonal direction. The images were acquired using a Canon 5D equipped with a fisheye lens leveled on a tripod at around breast height (1.3 m above the ground) near dawn or dusk. Two-corner classification was applied on the images obtained, and a combined Lang and Xiang clumping correction was used to estimate the LAI as outlined by Woodgate et al. (2015), Macfarlane (2011), and Leblanc et al. (2005), respectively.

Stem density was calculated as the number of trees per hectare based on the number of trees in each plot. Crown closure was measured by averaging five observations within each plot using a spherical crown densiometer (Forestry Suppliers, Inc., Jackson, USA). Crown diameter was calculated from the mean of the measurements in two directions. The stand height was measured using a Nikon Forestry 550 laser rangefinder. Both crown diameter and stand height were obtained by averaging the values of five randomly selected trees in each plot. Table 5.1 presents the summary statistics of the field leaf and canopy parameters measured.

**Table 5.1** Summary statistics of the leaf parameters and canopy parameters measured in the field (for 137 leaf samples and 26 sampling plots)

Parameter	Abbreviation	Unit	Minimum value	Maximum value	Mean
<b>(1) Leaf parameters</b>					
Equivalent water thickness	EWT	cm	0.0063	0.0337	0.017
Leaf mass per area	LMA	$\frac{\text{g}}{\text{cm}^2}$	0.0034	0.0291	0.014
Leaf nitrogen content	$N_{\text{area}}$	$\frac{\text{g}}{\text{cm}^2}$	1.43E-04	3.68E-04	2.78E-04
<b>(2) Canopy parameters</b>					
Leaf area index	LAI	$\frac{\text{m}^2}{\text{m}^2}$	2.85	5.14	3.61
Stem density	SD	$\text{ha}^{-1}$	222	1722	771
Stand height	H	m	8	38	23
Crown diameter	CD	m	1.65	15.45	5.4
Canopy closure	CC	%	77	91	82

#### 5.2.1.4 Calculation of plot-level leaf and canopy nitrogen content

Within each plot, the stand height, the diameter at breast height (DBH), and the number of trees from each dominant species were used to calculate the species fraction of foliar biomass using published allometric equations (Gower et al. 1993; Widlowski et al. 2003). The mean leaf nitrogen content per plot ( $N_{\text{plot}}$ ,  $\text{g}/\text{cm}^2$ , per leaf area) was calculated as the mean leaf nitrogen content for each species, weighted by the species leaf area fraction (Homolova et al. 2013)

$$N_{\text{plot}} = \sum_{i=1}^k n_i \text{fLAI}_i \quad (5.1)$$

where  $n_i$  represents the average leaf nitrogen content of species  $i$  within a plot,  $\text{fLAI}_i$  is the leaf area fraction of species  $i$  (in  $\text{g}/\text{g}$ ), and  $k$  is the number of tree species within a plot.

The species leaf area fraction was calculated by the species foliar biomass fraction and specific leaf area ( $\text{cm}^2/\text{g}$ , the inverse ration of LMA) (Martin et al. 2008)

$$fLAI_i = \frac{fBiomass_i SLA_i}{\sum_{i=1}^k fBiomass_i SLA_i} \quad (5.2)$$

where  $fBiomass_i$  is the foliar biomass fraction of species  $i$ , and  $SLA_i$  is the average specific leaf area for species  $i$  within a plot.

The canopy nitrogen content ( $g/m^2$ , gram per ground area) was calculated by the product of LAI and plot-level mean leaf nitrogen content.

#### 5.2.1.5 Spectral measurements for forest background

The background spectra were measured using an ASD FieldSpec-4 Pro spectrometer (Analytical Spectral devices, Inc., Boulder, CO, USA) coupled to a high density contact probe. No bare soil was found in the sampling plots, so the spectra from the background elements, such as bark, stem, dried leaves, understory and moss, were collected from a variety of representative plots.

### 5.2.2 Airborne hyperspectral data collection and processing

The hyperspectral data was obtained with the HySpex sensor by the German Aerospace Center (DLR) for the study area on 22 July 2013. The HySpex sensor-system consists of two imaging spectrometers with spectral ranges of 400–1000 nm (visible and near infrared, VNIR) and 1000–2500 nm (short-wave infrared), which record the solar radiance reflected from the Earth's surface. The HySpex sensor comprises 160 spectral channels and 256 channels, with spectral resolutions of 3.7 and 6 nm, and spatial resolutions of 1.65 m and 3.3 m, for VNIR and SWIR, respectively. The HySpex data were recorded between 9:00 and 11:00 a.m., at an average flying height of 3000 m above ground level. The data were collected in 19 image strips for the study area with overlaps of about 30% and each strip covering about  $1.1 \times 11$  km. The flight line was run in an almost N–S direction. Most of the image strips were acquired in clear weather conditions and none of the sample plots in the strips were covered by clouds.

The image strips were preprocessed according to a standard procedure detailed in Wang et al. (2016). The image reflectance values for each field plot were extracted from the preprocessed strips. A  $17 \times 17$  pixel window (i.e.  $27.2 \times 27.2$  m) centered on the center point of each plot was used to collect the spectra in VNIR from a sampling plot, while a  $9 \times 9$  pixel window (i.e.  $28.8 \times 28.8$  m) was selected for SWIR. For each window, the average spectrum to represent a plot was calculated. In total, 26 spectra were extracted over the 800–2500 nm, in which spectral region leaf nitrogen contributed most to spectra.

The HySpex image strips were first resampled to a spatial resolution of 30 m and then mosaicked to a single image. A forest mask was applied to the image to extract the forest areas and a forest map was derived from recently updated land cover data provided by the Bavarian Forest National Park.

### **5.2.3 Model description, parameterization and inversion**

#### **5.2.3.1 Leaf optical properties model: PROSPECT**

The PROSPECT leaf optical properties model was developed to simulate the leaf directional-hemispherical reflectance and transmittance over the optical domain from 400–2500 nm (Jacquemoud and Baret 1990). The model only requires a few input parameters, including leaf structure parameter ( $N_{\text{struc}}$ ), leaf chlorophyll content ( $C_{\text{ab}}$ ,  $\mu\text{g}/\text{cm}^2$ ), equivalent water thickness (EWT, cm) and leaf mass per area (LMA,  $\text{g}/\text{cm}^2$ ). The model was improved and recalibrated by Feret et al. (2008), resulting in the new versions PROSPECT-4 and PROSPECT-5. PROSPECT-5 is the same as PROSPECT-4 except for the separation of total leaf chlorophyll and total leaf carotenoids ( $C_{\text{cx}}$ ,  $\mu\text{g}/\text{cm}^2$ ) in the visible range of 400–750 nm. Recently PROSPECT-5 was used to estimate protein ( $C_{\text{p}}$ ,  $\text{g}/\text{cm}^2$ ) and cellulose + lignin content ( $C_{\text{cl}}$ ,  $\text{g}/\text{cm}^2$ ) in fresh leaves by simulating their effects on leaf reflectance and transmittance (Wang et al. 2015a). The revised PROSPECT-5 model considers  $N_{\text{struc}}$ ,  $C_{\text{ab}}$ ,  $C_{\text{cx}}$ ,  $C_{\text{w}}$ ,  $C_{\text{p}}$ , and  $C_{\text{cl}}$  as input parameters, please see Wang et al. (2015a) for more details.

PROSPECT was originally developed for broadleaf species (Jacquemoud and Baret 1990) and it performed well on broadleaves (Feret et al. 2008; Jacquemoud et al. 1996; Jacquemoud et al. 2009; Wang et al. 2015b). Previous studies also proved that the model provided a reasonable description of needle optical properties (Ali et al. 2016c; Hernandez-Clemente et al. 2014; Jacquemoud et al. 2009; Laurent et al. 2011; Moorthy et al. 2008; Zarco-Tejada et al. 2004). The revised PROSPECT-5 described by Wang et al. (2015a) was based on broadleaf species found in the publicly available LOPEX dataset (Hosgood et al. 1995), although its applicability to conifer species was not tested due to the lack of a proper dataset.

### 5.2.3.2 Canopy reflectance model: the invertible forest reflectance model (INFORM)

The invertible forest reflectance model (INFORM) was developed to simulate the bi-directional reflectance in forest stands by combining the forest light interaction model (FLIM) (Rosema et al. 1992), Scattering by Arbitrary Inclined Leaves (SAILH), and PROSPECT. INFORM is one of the hybrid models which provides a trade-off between detailed characterization of the canopy structure and model invertibility. The INFORM model accounts for the 1-dimensional, turbid medium radiative-transfer within the crowns and the 3-dimensional effects, such as crown-created shadows, the hotspot and the clumping of leaves in crowns (Schlerf and Atzberger 2006, 2012). The model is an invertible model and has been used for estimating biophysical and biochemical forest parameters such as LAI, fAPAR, specific leaf area, and leaf dry matter content in both broadleaf and conifer stands (Schlerf and Atzberger 2006, 2012; Yang et al. 2011). However, it has not yet been used for retrieving leaf nitrogen.

The previous submodel of PROSPECT built into INFORM was updated by the revised PROSPECT-5, described in Section 5.2.3.1.

### 5.2.3.3 Model parameterization

The ranges of input parameters and their fixed values are listed in Table 5.2.

For the revised PROSPECT-5, leaf chlorophyll and carotenoids content were held constant since these two parameters have negligible effects on leaf spectra for wavelengths longer than 800 nm. Leaf structure parameter has a small influence on canopy reflectance (Jacquemoud 1993; Xiao et al. 2014), thus this was fixed at 1.5 based on previous studies (Ali et al. 2016b; Ali et al. 2016c). Leaf nitrogen content was retrieved from laboratory measurements and the leaf protein content was calculated as 6.25 times of leaf nitrogen content (Barton 1987). The ranges of leaf water and protein content were determined from field measurements. Leaf cellulose + lignin content varied from 0.00034 to 0.027 based on the range of the difference between leaf mass per area and leaf protein content (by assuming that the leaf dry mass is composed of protein and cellulose + lignin).

The ranges of the INFORM input parameters, including stem density, crown diameter and stand height, were set based on prior information from the field (Table 5.2). Single-tree LAI was calculated as the ratio of LAI and canopy closure as demonstrated in Schlerf and Atzberger (2006), which ranged from 2.5 to 8. The range of average leaf inclination angle (ALA) was set according to Ali et al.

(2016b). The understory LAI was fixed at 0.1. The soil brightness parameter (scale) was set to 0.5, as suggested in Schlerf and Atzberger (2012). The ratio of diffuse to total incident radiation (SKYL) was set to 0.1, as suggested in the literature (Darvishzadeh et al. 2008; Schlerf and Atzberger 2006). The ranges of measurement geometry, including zenith angle of solar (tts), observation zenith angle (tto), and relative azimuth angle ( $\psi$ ) were acquired from the HySpex image acquisition campaign. The HySpex observation zenith angle was 0, the sun zenith angle ranged from 28.7 to 38.7 based on the flight time period, and the azimuth angle varied from 126.3 to 182.

We used the average reflectance spectrum that was measured for understory and forest floor in representative plots (see Section 5.2.1.5) as a fixed background reflectance in the model.

**Table 5.2** The input parameters and their ranges used for generating the look-up table, using forward modes in the INFORM.

Parameter	Abbreviation	Unit	Minimum value	Maximum value
Leaf structure index	$N$	-	1.5	1.5
Leaf chlorophyll content	$C_{ab}$	$\mu\text{g}/\text{cm}^2$	40	40
Leaf carotenoids content	$C_{cx}$	$\mu\text{g}/\text{cm}^2$	10	10
Equivalent water thickness	EWT	cm	0.0063	0.0337
Leaf protein content	$C_p$	$\text{g}/\text{cm}^2$	0.0009	0.0024
Leaf cellulose+lignin content	$C_{cl}$	$\text{g}/\text{cm}^2$	0.0025	0.027
Single-tree leaf area index	$\text{LAI}_s$	$\text{m}^2/\text{m}^2$	2	8
Understory leaf area index	$\text{LAI}_u$	$\text{m}^2/\text{m}^2$	0.1	0.1
Average leaf inclination angle	ALA	degree	40	60
Soil brightness parameter	scale	-	0.5	0.5
Stem density	SD	$\text{ha}^{-1}$	200	1800
Stand height	H	m	8	38
Crown diameter	CD	m	3	11
Sun zenith angle	$t_s$	degree	28.7	38.7
Observation zenith angle	$t_o$	degree	0	0
Azimuth angle	$\psi$	degree	126.3	182
Fraction of diffuse radiation	Skyl	-	0.1	0.1



#### 5.2.3.4 Global sensitivity analysis

Sensitivity analysis helps to identify the contribution of variation in input parameters to the variability in the output canopy reflectance. There are two types of sensitivity analyses, i.e., local and global sensitivity analyses. Local sensitivity analysis provides information on how the variation of each input parameter individually explains the variation in the model output and ignores the interactions between model parameters, while global sensitivity analysis provides information of how the variation of each input parameter individually, and their interactions with each other, account for the variation of model output (Asner 1998; Saltelli 1999). Global sensitivity analysis includes the simultaneous variations of model parameters (Bowyer and Danson 2004), we therefore adopted a global sensitivity analysis in this study. The global sensitivity analysis was performed for INFORM using the ranges of input parameters presented in Table 5.2. A Matlab software tool (GSAT) (Cannavó 2012) was applied to perform the global sensitivity analysis; see Wang et al. (2015a) for more details.

#### 5.2.3.5 The look-up table inversion

There are a number of inversion approaches, such as iterative optimization, look-up table (LUT), and neural network (Kimes et al. 2000). The look-up table (LUT) inversion approach was chosen in this study, since it is a conceptually simple technique, can be easily implemented, and yields similar results to the alternatives (Combal et al. 2002; Pragnere et al. 1999). The advantages of LUT inversion are computation efficiency, a guarantee to find a global minimum, as well as the ability to apply ecological constraints during the inversion procedure (Combal et al. 2002; Houborg et al. 2009; Jurdao et al. 2013; Yebra et al. 2008).

Prior to inversion, a LUT was built via forward modeling using different combinations of input parameters covering their prescribed range of variation (Table 5.2). A sufficiently large LUT is needed to ensure high accuracy for the estimated parameters, as suggested by (Darvishzadeh et al. 2008). In this study, 200,000 parameter combinations (uniform distributions) were randomly generated from the forward modeling of INFORM. To reduce unrealistic combinations of input parameters in the look-up table, we applied two ecological rules obtained from field measurements to filter out some of the simulations. The first filter utilized the empirical relationship between LAI and equivalent water thickness ( $LAI = 106.43 \cdot EWT + 2.57$ ,  $R^2 = 0.64$ ). The cases that exceeded 10% of the range derived from the maximum or minimum residue of the regression fitting were excluded. The threshold of 10% was suggested in Yebra et al. (2008).

We also filtered the look-up table by selecting the cases with canopy closure between 0.1 and 0.95, because values beyond this range were rarely found in our study area. After applying the filtering criterions, 81,857 records were retained in the look-up table.

Then, a set of parameters was identified by searching for the best fit between measured spectra and the modeled spectra of the LUT by minimizing the

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (R_{mes}(\lambda) - R_{LUT}(\lambda))^2}{n}} \quad (5.3)$$

where  $R_{mes}(\lambda)$  is the measured reflectance at wavelength  $\lambda$ ,  $R_{LUT}(\lambda)$  is the reflectance modeled by INFORM and stored in the LUT at wavelength  $\lambda$ , and  $n$  is the number of wavelengths. To overcome the ill-posed problem, previous studies suggested using the mean or median of the parameters corresponding to the first 100 best matches instead of those from the best fit (Darvishzadeh et al. 2008; Schlerf and Atzberger 2012). Therefore, 100 best matched spectra were identified, and the mean of their 100 sets of corresponding parameters were calculated as the estimates of the targeted parameters.

Previous studies have demonstrated that the use of spectral subsets rather than the full wavelengths could provide equally or more accurate estimations through model inversion (Darvishzadeh et al. 2011; Darvishzadeh et al. 2008; Wang et al. 2015a; Weiss et al. 2000). Extra bands may add some noise instead of adding useful information on model inversion, which is due to the model uncertainties at certain wavelengths as well as the uncertainties in measurements (Weiss et al. 2000). We therefore performed the inversion over the following spectral subsets presented in Table 5.3: (1) full available wavelength range; (2) the spectral regions excluding water absorption bands; (3) the spectral regions where protein contributed relatively more to reflectance according to the global sensitivity analysis; and (4) protein absorption wavelengths based on the literature (Curran 1989; Fourty et al. 1996).

**Table 5.3** Selected spectral subsets for model inversion to estimate leaf and canopy nitrogen content.

Spectral subset	Wavelengths (nm)
1	800-2500
2	800-1350, 1550-1750, 2000-2400
3	1550-1800, 2100-2300
4	1020, 1510, 1730, 1980, 2060, 2130, 2180, 2240, 2300

#### 5.2.4 Validation

For both leaf and canopy level analysis, the leaf nitrogen content was calculated as the estimated leaf protein content divided by 6.25. The accuracy of the retrieved leaf nitrogen content, LAI and canopy nitrogen content via INFORM inversion was evaluated by the coefficient of determination ( $R^2$ ), the root mean square error (RMSE), and the normalized RMSE (NRMSE = RMSE/range) between the estimated and measured values.

### 5.3 Results

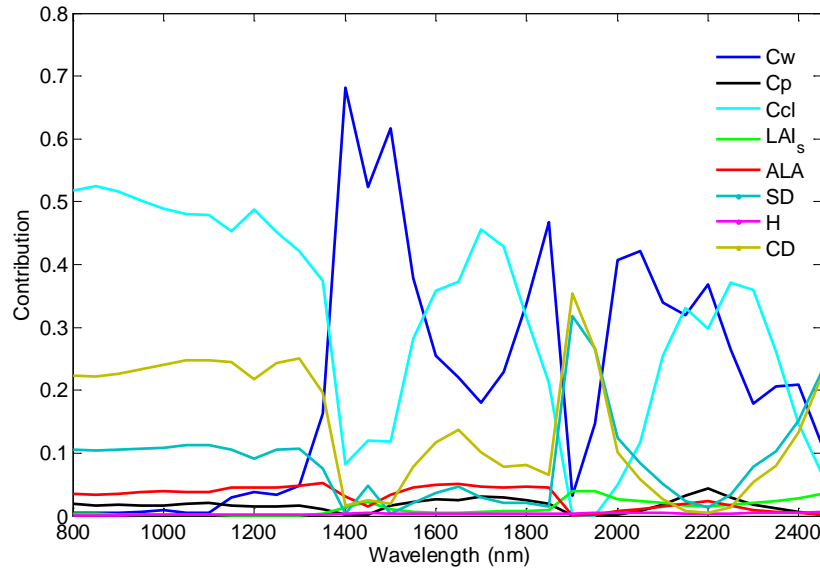
#### 5.3.1 Characteristics of leaf and canopy parameters

Table 5.1 gives the statistical characteristics of leaf and canopy parameters of the 26 sampling plots. The leaf nitrogen content varied from 1.43E-04 to 3.68E-04 g/cm<sup>2</sup>, with a mean of 2.78E-04 g/cm<sup>2</sup>. The mean LAI of all plots was 3.61, with a minimum value of 2.85 and a maximum value of 5.14. The mean stem density, tree height, crown diameter and canopy closure of the sampling plots were 771 ha<sup>-1</sup>, 23 m, 5.4 m, and 82%, respectively.

#### 5.3.2 Global sensitivity analysis

Fig. 5.2 shows the FAST first-order sensitivity coefficients of the input parameters to canopy reflectance for a global sensitivity analysis with INFORM model. Cellulose + lignin contributed most to canopy reflectance (around 50%) in the near infrared spectral region of 800 – 1300 nm, followed by crown diameter (20%) and stem density index (10%). In the wavelengths longer than 1300 nm, water dominated the canopy reflectance, contributing up to 60% of the explained variance. In the weak absorption wavelengths of water, cellulose + lignin contributed most to canopy reflectance over the spectral intervals 1500 – 1850 nm and 2100 – 2300 nm. Stem density and crown diameter had pronounced

effects on canopy reflectance in the spectral intervals of 1850 – 2150 nm and 2300 – 2500 nm. In comparison, single-tree LAI, the average leaf inclination angle, and stand height had less influence on canopy reflectance over the interval 800 – 2500 nm. Protein had a small but noticeable impact on canopy reflectance (around 2 – 4%) in the spectral regions with weaker water absorption.



**Figure 5.2** Results of FAST first-order sensitivity coefficients and interactions to canopy reflectance for the global sensitivity analysis with INFORM.

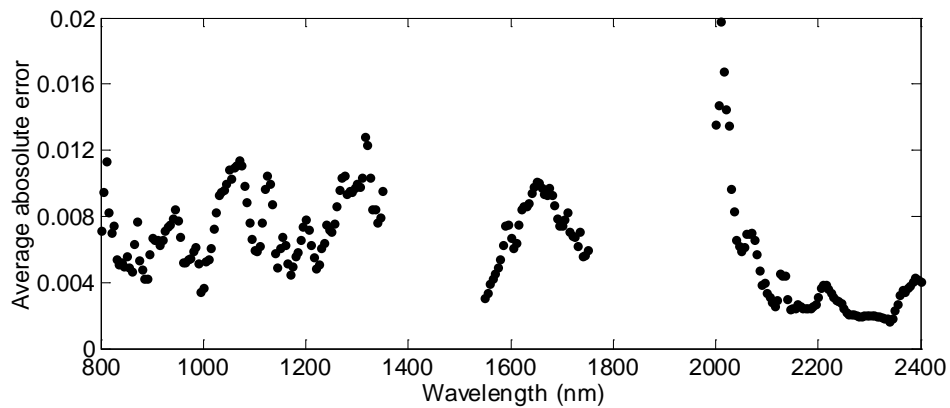
### 5.3.3 Validation of estimated leaf and canopy nitrogen content using INFORM inversion

Table 5.4 gives the validation leaf nitrogen content, LAI and canopy nitrogen content estimated from INFORM. Leaf nitrogen content was estimated with a relatively low accuracy from the INFORM model inversion ( $R^2 = 0.38$ , RMSE = 0.000040, NRMSE = 0.20). LAI was retrieved with a  $R^2$  of 0.54 and RMSE of 0.68. More accurate estimation was obtained for canopy nitrogen content ( $R^2 = 0.61$ , RMSE = 2.40, NRMSE = 0.22), which can be mainly attributed to the higher accuracy of the LAI estimation. For all three parameters, the most accurate estimates were provided by the mean of the best 100 cases found in the LUT. The spectral subsets 1, 2 and 4 generated similar accuracies, while subset 3 provided poorer estimates (Table 5.4). Using the wavelengths adopted for subset 2, an accurate estimation of canopy reflectance was obtained through parameterization of the models, as demonstrated by the low average absolute error (AAE)

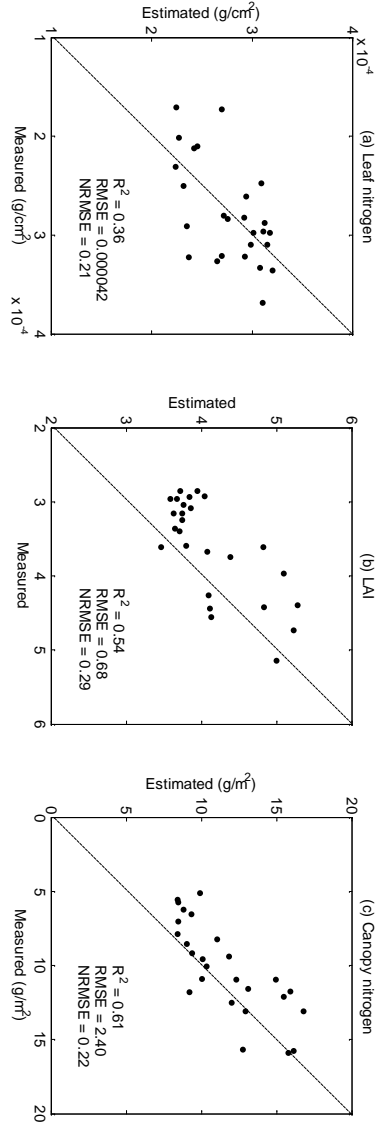
(Darvishzadeh et al. 2008) between measured and best-fit reflectance spectra (AAE <0.02, Fig. 5.3).

**Table 5.4** The  $R^2$ , RMSE, NRMSE between measured and estimated leaf nitrogen content, LAI and canopy nitrogen content obtained using inversion of INFORM.

Spectral sampling set	Statistical parameter	Leaf nitrogen content (g/cm <sup>2</sup> )			LAI			Canopy nitrogen content (g/m <sup>2</sup> )		
		$R^2$	RMSE	NRMSE	$R^2$	RMSE	NRMSE	$R^2$	RMSE	NRMSE
Subset 1	Best fitting spectra	0.24	0.000069	0.35	0.23	0.88	0.39	0.53	3.65	0.34
	Mean of 10	0.25	0.000054	0.28	0.57	0.81	0.36	0.54	3.55	0.33
	Mean of 100	0.36	0.000042	0.21	0.67	0.82	0.36	0.67	2.89	0.27
Subset 2	Best fitting spectra	0.04	0.000085	0.43	0.20	0.79	0.35	0.29	3.65	0.34
	Mean of 10	0.31	0.000048	0.24	0.40	0.67	0.29	0.44	3.09	0.29
	Mean of 100	<b>0.36</b>	<b>0.000042</b>	<b>0.21</b>	<b>0.54</b>	<b>0.68</b>	<b>0.30</b>	<b>0.61</b>	<b>2.40</b>	<b>0.22</b>
Subset 3	Best fitting spectra	0.07	0.000082	0.41	0.43	0.80	0.35	0.23	4.22	0.39
	Mean of 10	0.11	0.000053	0.27	0.40	0.60	0.26	0.33	2.93	0.27
	Mean of 100	0.23	0.000045	0.23	0.41	0.55	0.24	0.38	2.50	0.23
Subset 4	Best fitting spectra	0.08	0.000091	0.46	0.18	0.81	0.35	0.19	5.34	0.49
	Mean of 10	0.24	0.000056	0.29	0.54	0.67	0.29	0.48	3.42	0.32
	Mean of 100	0.38	0.000040	0.20	0.57	0.67	0.29	0.60	2.53	0.23



**Figure 5.3** The average absolute error (AAE) between measured and best-fit reflectance spectra as a function of wavelengths. The AAE has been calculated from the 26 measured canopy spectra against the best fitting look-up table (LUT) spectra for spectral subset 2 from INFORM.



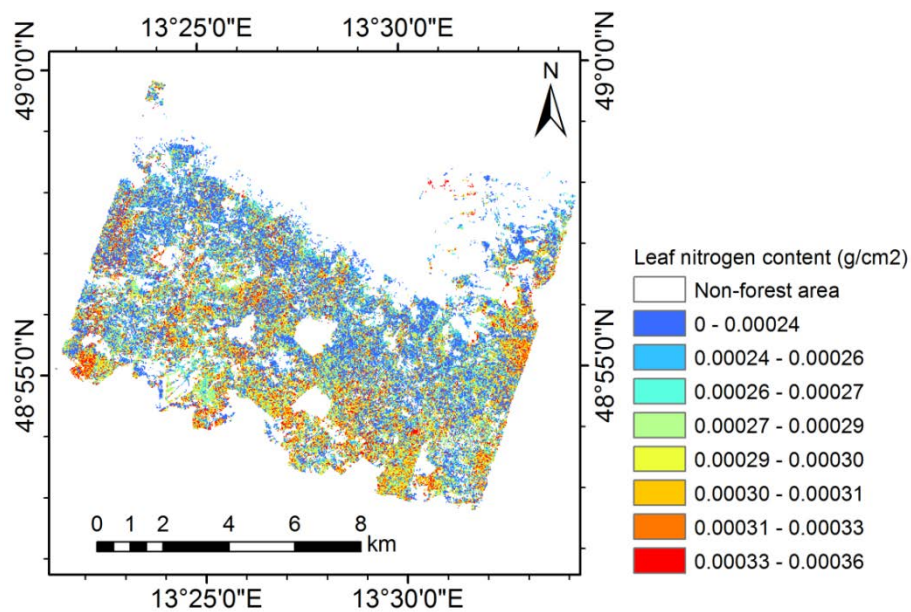
**Figure 5.4** Measured versus estimated parameters: (a) leaf nitrogen content (g/cm<sup>2</sup>) (b) LAI, and (c) canopy nitrogen content (g/m<sup>2</sup>) from measured spectra using spectral subset 2 by INFORM inversion.

#### 5.3.4 Mapping of leaf and canopy nitrogen content from airborne hyperspectral imagery

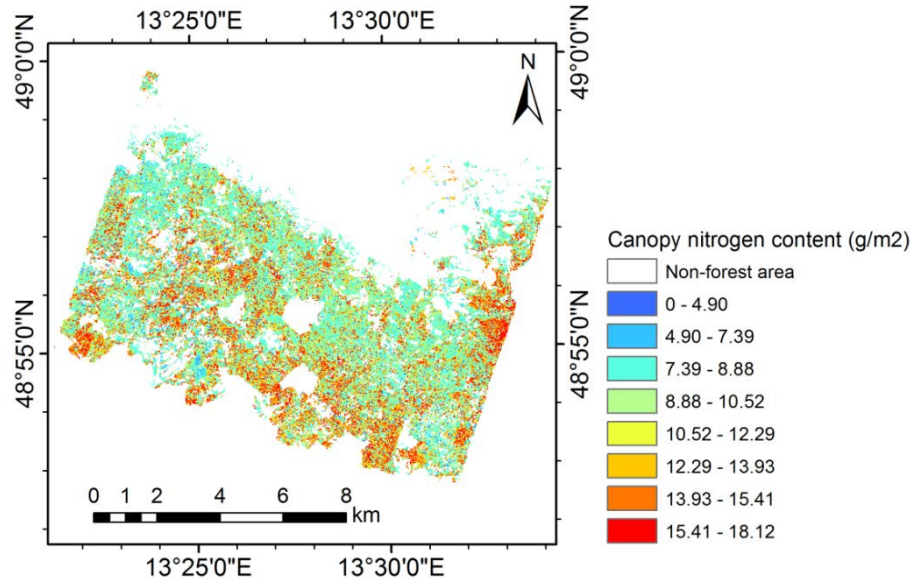
The masked HySpex image of forest areas (see Section 5.2.2) was used as input to the INFORM inversion, resulting in a map of leaf nitrogen content (Fig. 5.5).

LAI is a critical forest structural parameter used to scale up the leaf-level biochemical parameters to canopy level. LAI cannot be directly obtained by inversion of INFORM, but is calculated as the product of retrieved single-tree

LAI and canopy closure; canopy closure was calculated using the retrieved stem density and crown diameter (Schlerf and Atzberger 2006). Canopy nitrogen content per unit ground surface area is defined as the product of leaf nitrogen content per unit leaf area and LAI. Canopy nitrogen content was mapped (Fig. 5.6) based on the INFORM leaf nitrogen content map and LAI map. The spatial variation of leaf nitrogen content in the generated map corresponds well with the distribution of broadleaf, needle leaf and mixed forest observed during the fieldwork.



**Figure 5.5** Leaf nitrogen content (g/cm<sup>2</sup>) estimated using the HySpex airborne image data and inversion of the INFORM mode in the Bavarian Forest National Park, Germany.



**Figure 5.6** Canopy nitrogen content (g/m<sup>2</sup>) calculated using the HySpex airborne image data and inversion of the INFORM model in the Bavarian Forest National Park, Germany.

## 5.4 Discussion

This study has confirmed the feasibility of retrieving leaf nitrogen content at the canopy level by coupling PROSPECT-5, a leaf radiative transfer model, and INFORM, a canopy reflectance model. We recalibrated the PROSPECT-5 model by incorporating the effects of leaf protein and cellulose + lignin content, which enabled us to estimate the leaf nitrogen content from fresh leaf spectra (Wang et al. 2015a). When we linked the revised PROSPECT-5 with a canopy reflectance model such as INFORM, we could retrieve leaf and canopy nitrogen content at the canopy level using airborne hyperspectral imagery.

Prediction of leaf biochemistry at the canopy level faces a number of challenges, such as the confounding factors of canopy structure, illumination/viewing geometry and background (Asner 1998; Yoder and Pettigrew-Crosby 1995; Zarco-Tejada et al. 2001). Previous studies have demonstrated that lower accuracies were obtained for leaf biochemistry compared with canopy biochemistry (Darvishzadeh et al. 2008; Liang et al. 2016; Omari et al. 2013; Si et al. 2012). In these studies, an  $R^2$  of 0.14 – 0.40 was obtained for leaf chlorophyll and an  $R^2$  of 0.60 – 0.80 for canopy chlorophyll in grasslands and trembling aspen forest when coupled leaf-canopy radiative transfer models were inverted. Another study reported an  $R^2$  of 0.47 for retrieved leaf chlorophyll from the canopy spectra



for black spruce forest (Zhang et al. 2008). Given the small quantity and small range of nitrogen in leaves, the leaf nitrogen content was retrieved at a reasonable accuracy in this study ( $R^2 = 0.38$ , NRMSE = 0.20). It should be noted that the spectral wavelengths from 400 – 799 nm were not incorporated in this study, because the PROSPECT-5 model was only calibrated in the spectral interval of 800 – 2500 nm. The existing correlation between nitrogen and chlorophyll means a higher accuracy for nitrogen estimation may be expected if red-edge region (680 – 780 nm) information is included.

Despite the above confounding factors, our results demonstrated that leaf nitrogen content can be detected at the canopy level using canopy reflectance. In previous studies which focused on empirical models, increased predictability of leaf biochemistry was observed when up-scaling to the canopy level (Asner and Martin 2008b; Asner et al. 2011). One reason for the feasibility of estimating leaf nitrogen content at the canopy level is the strong, multiple scattering effects in forests, which can enhance the leaf biochemical signals by up to a factor of two (Baret et al. 1994). In other words, due to the relatively high LAI, the light scattering and absorption enhances the spectroscopic difference among canopies and, hence, leaf biochemistry retrieval (Asner and Martin 2008b).

The second explanation is that the parameterization of the INFORM model is probably quite realistic. The INFORM model includes parameters for characterizing the canopy geometrical structure, and incorporates the forest light interaction model (FLIM) model to account for 3-dimensional effects such as shadows, the hotspot, and clumping of leaves in crowns (Schlerf and Atzberger 2006, 2012). These enable the disassociation of the remotely sensed signal from canopy structural effects and leaf biochemical properties.

Global sensitivity analysis showed that leaf water contributed most to the variation of canopy reflectance, which is consistent with previous studies (Jacquemoud et al. 2009; Xiao et al. 2014). Leaf cellulose + lignin exerted a great influence on canopy reflectance, particularly in the NIR spectral regions (800 - 1300 nm). This indicated that the impact of the leaf constituents associated with internal leaf structure is transferred to the canopy level. In terms of the canopy structural parameters, we found stand density and crown diameter were the dominant factors affecting canopy reflectance, while less influence was found from single-tree LAI and the average leaf inclination angle. This is in agreement with a recent study focusing on canopy structural effects on retrieving specific leaf area and leaf dry matter content from remotely sensed data (Ali et al. 2016b).

In the INFORM model, canopy LAI is correlated to single-tree LAI, stem density and crown diameter, with the contribution of canopy LAI being mostly compensated by stem density and crown diameter. Therefore, a larger influence of canopy LAI on canopy reflectance is expected if single-tree LAI is converted to canopy LAI through stem density and crown diameter. Leaf protein had a small but noticeable effect on the variation of canopy reflectance (e.g. 800 – 1300 nm, 1600 – 1800 nm, 2100 – 2300 nm) as simulated using INFORM.

RTMs inherently embody the ill-posed inverse problem (Combal et al. 2002). Firstly, different combinations of input parameter may generate similar spectra. Secondly, measurement and model uncertainties will lead to inconsistencies during model inversion. Regularization techniques have been proposed to minimize the ill-posed inverse problem, such as using prior information (Combal et al. 2002), spectral subsets (Darvishzadeh et al. 2008; Wang et al. 2015a; Weiss et al. 2000), ecological constraints (Jurdao et al. 2013; Yebra and Chuvieco 2009), and a spatial regularization approach (Houborg et al. 2009; Houborg et al. 2015).

In this study, we used the prior information from field measurements, which constrains the input parameters and avoids unrealistic solutions. Ecological rules were applied to the look-up table to reduce unrealistic combinations of input parameters. Instead of adopting the best case LUT, the mean of the first 100 best cases was selected as the solution. The latter provided more robust results for the retrieved parameters (Table 5.4). The spectral subsets 1, 2 and 4 generated similar accurate estimations, while subset 3 provided a poorer estimate (Table 5.4). This indicated that the spectral subset had less influence on the accuracy of the estimation (Darvishzadeh et al. 2008) compared to the other two regularization techniques mentioned above, such as using ecological constraints and multiple solutions. In all cases, leaf water content was estimated with a good accuracy ( $R^2 = 0.63 - 0.76$ , results not shown). Due to the masking effect of water, water removal techniques are required in further research (Wang et al. 2015a), such as integration with continuum removal (Clark and Roush 1984; Malenovsky et al. 2013) as well as with wavelet transforms (Banskota et al. 2013).

Figure 5 demonstrated that the spatial pattern of leaf nitrogen content in forest maps corresponds well with the distribution of broadleaf, needle leaf and mixed forest. The areas with higher nitrogen content are in good agreement with the distribution of needle leaf forest, which is mainly found in the southern valley areas, while low nitrogen areas are consistent with the broadleaf forest located in the north at a higher elevation. This was also confirmed by our observations

during the fieldwork (Wang et al. 2015b). We calculated the mean and standard deviation of leaf nitrogen content for all image pixels in the forest area. The mean and standard deviation of INFORM leaf nitrogen were  $2.5\text{E-}04$  and  $9.1\text{E-}05$ , respectively, which is close to values of sampled plots measured in the field ( $2.8\text{E-}04$  and  $5.2\text{E-}05$ ). The canopy nitrogen content per unit ground surface area ranged from  $4.90 - 18.12 \text{ g/m}^2$  for forest pixels (Fig. 5.6). The canopy nitrogen map shows a similar pattern to the leaf nitrogen map, with the variance of canopy nitrogen across different plant functional types being driven more by the values of leaf nitrogen than LAI.

## 5.5 Conclusions

We have evaluated the feasibility of combining leaf and canopy radiative transfer models to retrieve leaf and canopy nitrogen content using airborne hyperspectral measurements. Leaf protein had a small but noticeable effect on the canopy reflectance. The look-up table detailed the canopy information and viewing geometry, which enabled the separation of canopy structural effects from leaf optical properties. Canopy nitrogen content was retrieved at a higher accuracy than leaf nitrogen content, which can be attributed to having good estimates of LAI. Inversion techniques, such as using prior information, ecological constraints, spectral subsets, and statistical parameters of a certain number of best solutions further improved our estimation accuracy. The most accurate estimation of leaf and canopy nitrogen was obtained when using spectral subset 2 and the mean of the first 100 cases. The leaf nitrogen and canopy nitrogen maps were generated by applying the inversion procedure to the whole hyperspectral imagery, and the spatial variation corresponded well with the distribution of plant functional types.

The robustness and transferability of radiative transfer models suggests that the approaches proposed in this study may be transferred to other sites with different natural and environmental conditions. More accurate estimation may result from using the recalibrated PROSPECT specifically for conifer needle species, as well as incorporating the spectral interval of  $400 - 799 \text{ nm}$ . With the aid of light detection and ranging (LiDar), canopy structural parameters that are inputs to the canopy model can be obtained and used as constraints in the model inversion (Asner et al. 2015; Combal et al. 2002; Gokkaya et al. 2015; Niemann et al. 2012). Larger scale maps of foliar nitrogen could be generated for modeling ecosystems and assessing biodiversity if hyperspectral satellites such as EnMAP (Guanter et al. 2015) and HypsIRI ((NRC) 2007) become operational.



## **Chapter 6**

### **Synthesis: Hyperspectral remote sensing of foliar nitrogen**

## **6.1 Introduction**

Foliar nitrogen is a critical parameter in leaf photosynthesis, respiration and transpiration (Evans 1989; Field and Mooney 1986; Reich 2012; Reich et al. 1998; Reich et al. 2006), which relates to canopy and stand-level traits such as light use efficiency, wood growth and net primary production (Ollinger and Smith 2005; Scott Green et al. 2003; Smith et al. 2002). Nitrogen is regarded as a limiting nutrient for plant growth (Heimann and Reichstein 2008; LeBauer and Treseder 2008), and it is one of the key plant traits driving stand-level forest productivity (Reich 2012). As a consequence, foliar nitrogen content has recently been proposed as one of the key essential biodiversity variables (EBVs) for satellite monitoring of progress towards the Aichi Biodiversity Targets (Pereira et al. 2013; Skidmore et al. 2015). Accurate estimation of nitrogen can be an important input parameter of ecosystem process models (Ollinger and Smith 2005; Plummer 2000; Smith et al. 2002).

Hyperspectral data provides a time-efficient and cost-effective solution to estimate foliar nitrogen compared to traditional destructive sampling methods. Previous studies on nitrogen estimates range from using spectra of ground leaf powder to dry leaves and to fresh leaves, also cause variation at leaf and canopy level (Martin and Aber 1997; Smith et al. 2002; Wessman et al. 1988). Empirical approaches have been the dominant method when estimating nitrogen, and good accuracies have been achieved, though any empirical relationship so established is hard to be transferred across species, sites and sensors (LaCapra et al. 1996; Martin and Aber 1997). Furthermore, wavelengths selected in different studies can be somewhat inconsistent and deviate from known absorption wavelengths (Curran et al. 2001; Huang et al. 2004; Jacquemoud et al. 1996). Physically-based approach, or radiative transfer models simulate the process of radiation transfer based on physical laws, which is more robust and transferable (Darvishzadeh et al. 2011; Jacquemoud and Baret 1990; Schlerf and Atzberger 2006). One of the prominent leaf radiative transfer models, the leaf PROSPECT (Jacquemoud and Baret 1990), tried to incorporate nitrogen in the model, but the idea was abandoned due to inconsistent inversion results (Jacquemoud et al. 1996). Nitrogen has been considered to be impossible to be retrieved from fresh spectra by inversion of a leaf RTM model (Jacquemoud et al. 1996), which remains an unsolved problem.

A leaf radiative transfer model incorporating nitrogen as an input parameter is the foundation of upscaling to canopy level. When estimating nitrogen at the canopy

level, a number of factors impede the detection of the nitrogen signal, such as canopy structure, illumination/viewing geometry and background (Asner 1998; Ross 1981; Zarco-Tejada et al. 2001). A debate about the role of canopy structure in nitrogen estimation started after a finding was reported that a positive relationship was found between nitrogen and canopy NIR reflectance (Ollinger et al. 2008). Knyazikhin et al. (2013a) argued that the positive relationship can be attributed to the influence of canopy structure between broadleaf and coniferous stands on canopy NIR reflectance, and is not driven by nitrogen variation. Knyazikhin et al. (2013a) proposed a simple approach to correct the influence of canopy structure by normalizing canopy BRF by a structural parameter, DASF, which is derived based on physical laws. After correcting the canopy structural effects, the relationship between nitrogen and canopy NIR reflectance turned to be negative. In other words, the results of Knyazikhin et al. (2013a) suggested that nitrogen may not be accurately estimated using remote sensing. However, the research was limited to wavelengths between 400 - 850 nm, and the wavelengths longer than 850 nm was not explored, a region in which most of the nitrogen absorption features are located (Curran 1989; Fourty et al. 1996; Kokaly et al. 2009). The coupling of leaf and canopy radiative transfer models provide another way of characterizing leaf biochemicals, leaf optical properties, canopy structure, illumination / viewing geometry and background information. Nitrogen estimates can be obtained at the canopy level by inversion of such coupled leaf and canopy radiative transfer models.

The research in this thesis initially evaluated the feasibility of retrieving nitrogen by inversion of a leaf radiative transfer model PROSPECT. Chapter 2 investigated the indirect estimation of nitrogen by inversion of leaf PROSPECT through links between nitrogen and leaf traits, and proved that nitrogen can be estimated at a moderate accuracy by inversion of the leaf PROSPECT model and the empirical relationship between nitrogen and leaf traits (e.g. leaf water, leaf mass per area and chlorophyll). Chapter 3 recalibrated the specific absorption coefficients of protein and cellulose + lignin, and proved that the recalibrated model can accurately simulate leaf reflectance and transmittance. Nitrogen can be estimated at a moderate accuracy by inversion of the recalibrated model. Then research moved on to the canopy level with emphasis on addressing canopy structural effects on nitrogen estimation. Chapter 4 evaluated the feasibility of nitrogen retrieval from the canopy BRF in the full wavelengths between 425 – 2500 nm corrected for canopy structure effects. The results showed that nitrogen can still be detected from canopy BRF after correcting for canopy structure, and

continuum wavelet analysis was shown to be an efficient method of revealing the nitrogen absorption features. The selected wavelet features most sensitive to nitrogen variation were found to be correlated with water, dry matter and protein. Chapter 5 tested the capability of coupled leaf-canopy models in foliar nitrogen estimation. Global sensitivity analysis showed that the influence of nitrogen content on canopy reflectance is small but noticeable. Canopy nitrogen content was more accurately predicted than leaf nitrogen content. Maps of leaf and canopy nitrogen were generated for the study area, and spatial variations of nitrogen appeared consistent with ecological knowledge.

## **6.2 Leaf nitrogen content indirectly estimated by leaf traits derived from the PROSPECT model**

Estimating leaf nitrogen in fresh leaves, using inversion of a radiative transfer model such as PROSPECT, has been considered impossible since the 1990s (Jacquemoud et al. 1996). However, the model parameters (leaf chlorophyll content,  $CHL_{area}$ ,  $g/cm^2$ ; leaf mass per area, LMA,  $g/cm^2$ ; and equivalent water thickness, EWT,  $g/cm^2$ ), which are also recognized as common leaf traits, can be retrieved via PROSPECT model inversion with intermediate to good accuracy (Baret and Fourty 1997a; Colombo et al. 2008; Darvishzadeh et al. 2008; Feret et al. 2008; le Maire et al. 2004; Omari et al. 2013; Riano et al. 2005; Zhang et al. 2008), and have been identified as good indicators of nitrogen (Ellsworth and Reich 1993; Evans 1989; Field and Mooney 1986). There are only a few spectroscopic studies focusing on nitrogen links to leaf traits (i.e., LMA and EWT), although its correlation with chlorophyll has been widely considered (Homolova et al. 2013; le Maire et al. 2008).

Chapter 2 explored the relationship between mass-based versus area-based leaf nitrogen content and leaf traits, and applied it to retrieve leaf nitrogen content from fresh leaf spectra combined with physically based models. Chapter 2 demonstrated an indirect estimation of leaf nitrogen by using leaf traits retrieved from PROSPECT model inversion. Further, area-based nitrogen correlations with leaf traits were found to be higher than mass-based correlations (Table 6.1). Regression models were derived for area-based leaf nitrogen content using highly correlated leaf traits (LMA, EWT and  $CHL_{area}$ ) as independent variables. The empirical models and retrieved leaf traits were combined to estimate leaf nitrogen content. Our results indicated that EWT was retrieved with a higher accuracy than LMA or  $CHL_{area}$  when inverting PROSPECT. Area-based leaf nitrogen content was estimated more accurately by regression models using EWT as predictor than

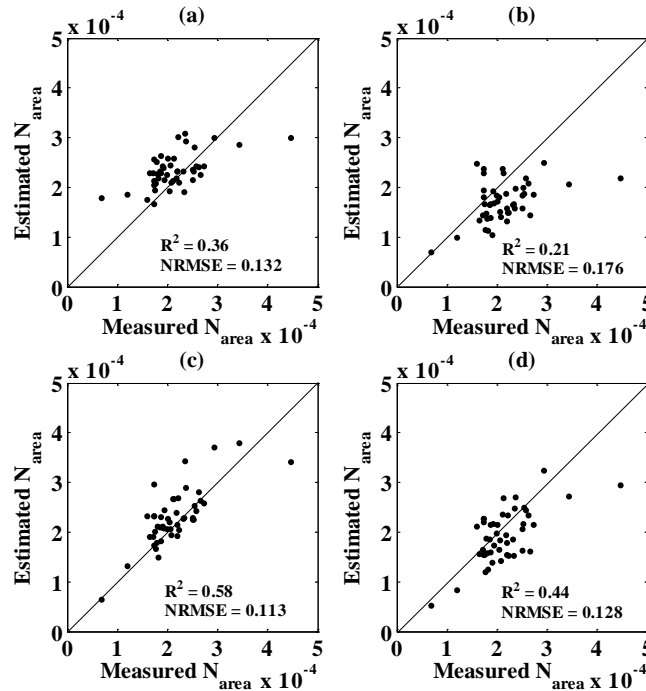


LMA or  $\text{CHL}_{\text{area}}$ . A linear regression (empirical) model using EWT as a predictor provided the most accurate estimation of nitrogen (Fig. 6.1). The combination of empirical and physically based models serves as a reliable method for estimating leaf nitrogen, although its transferability needs to be explored using other datasets.

**Table 6.1** Correlations between leaf nitrogen content and other leaf traits

	$\text{CHL}_{\text{area}}$	$\text{CHL}_{\text{mass}}$	LMA	LDMC	EWT	$\text{GWC}_t$
$\text{N}_{\text{area}}$	0.597**	-0.282	0.686**	0.066	0.841**	-0.066
$\text{N}_{\text{mass}}$	0.637**	0.555**	0.346*	0.663**	0.200	0.633**

\*\* correlations significant at  $p < 0.01$ , and \* correlations significant at  $p < 0.05$ . See Table 2.1 for the definitions of leaf traits.



**Figure 6.1** Comparison between measured and estimated  $\text{N}_{\text{area}}$  ( $\text{g}/\text{cm}^2$ ) using different regression models, coupled with their independent variables retrieved from the PROSPECT model inversion. The independent variables of each regression model were (a)  $\text{CHL}_{\text{area}}$ , (b) LMA, (c) EWT, (d) LMA and  $\text{CHL}_{\text{area}}$ .

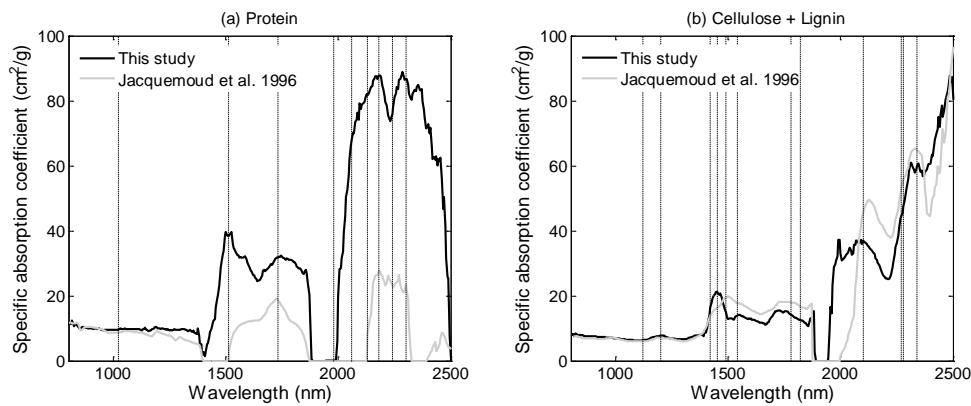
### 6.3 Retrieval of nitrogen by inversion of a leaf radiative transfer model PROSPECT

The inversion of PROSPECT for retrieving leaf nitrogen content in fresh leaves has remained a challenge and unresolved problem since the 1990s (Jacquemoud et al. 1996), although empirical methods achieved relatively good accuracies (Kokaly 2001; Kokaly et al. 2009; Kokaly and Clark 1999; Martin et al. 2008). Chapter 3 evaluated the feasibility of retrieving nitrogen through recalibrating PROSPECT-5 by incorporating the effects of leaf protein and cellulose + lignin on leaf reflectance and transmittance. To our knowledge, chapter 3 demonstrated, for the first time, that leaf nitrogen can be successfully estimated in fresh leaves at a moderate to good accuracy through inversion of a radiative transfer model, the leaf optical properties model PROSPECT-5 (Feret et al. 2008; Wang et al. 2015a).

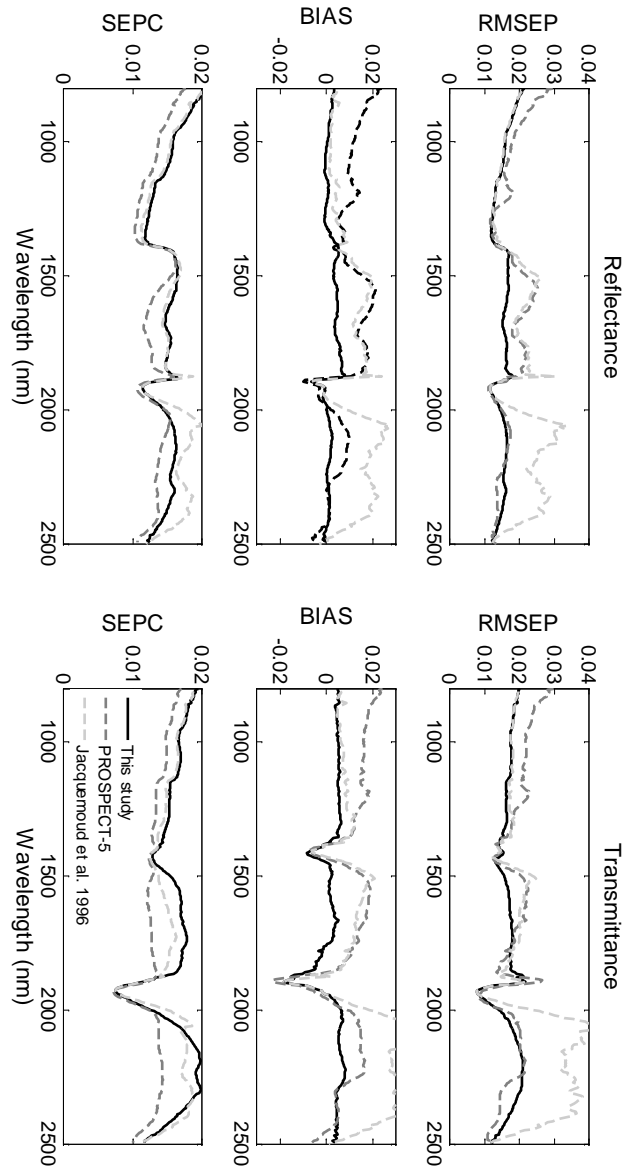
In Chapter 3, two challenges were addressed in the retrieval of leaf parameters (e.g. protein and cellulose + lignin) in fresh leaves by inversion of PROSPECT. The first issue is the determination of specific absorption coefficients of leaf biochemicals, which describes the absorption property of a leaf, together with the concentration of these biochemicals (Jacquemoud and Baret 1990). The absorption coefficients developed for PROSPECT-5 were recalibrated specifically for protein and cellulose + lignin using the new calibration algorithm proposed by Feret et al. (2008). Results showed great differences in the behavior of the coefficient of protein for fresh leaves over the ranges 1400 - 1800 nm and 2100 - 2500 nm, compared to those reported by Jacquemoud et al. (1996) (Fig. 6.2). The absorption peaks of protein reached approximately 80 cm<sup>2</sup>/g rather than their value of 20 cm<sup>2</sup>/g. The improvements of the coefficients were further confirmed when applying them in the PROSPECT model for forward simulation of leaf reflectance and transmittance, which can be largely attributed to the new calibration algorithm proposed by Feret et al. (2008). The leaf optical properties for fresh leaves were more accurately modelled, especially over the ranges 1400 - 1800 nm and 2100 - 2500 nm. The recalibrated coefficients were robustly estimated across species and showed a greater accuracy when reconstructing leaf reflectance and transmittance compared with a previous study (Jacquemoud et al. 1996) (Fig. 6.3). Absorption peaks in coefficients correspond well with known protein and cellulose + lignin absorption features regardless of the leaf water status (Fig. 6.2).

The “ill-posed” problem when inverting radiative transfer models is the second challenge in the retrieval of leaf protein and cellulose + lignin. Different combinations of leaf parameters can generate similar leaf spectra. And leaf parameters which make a greater contribution to leaf reflectance, such as leaf

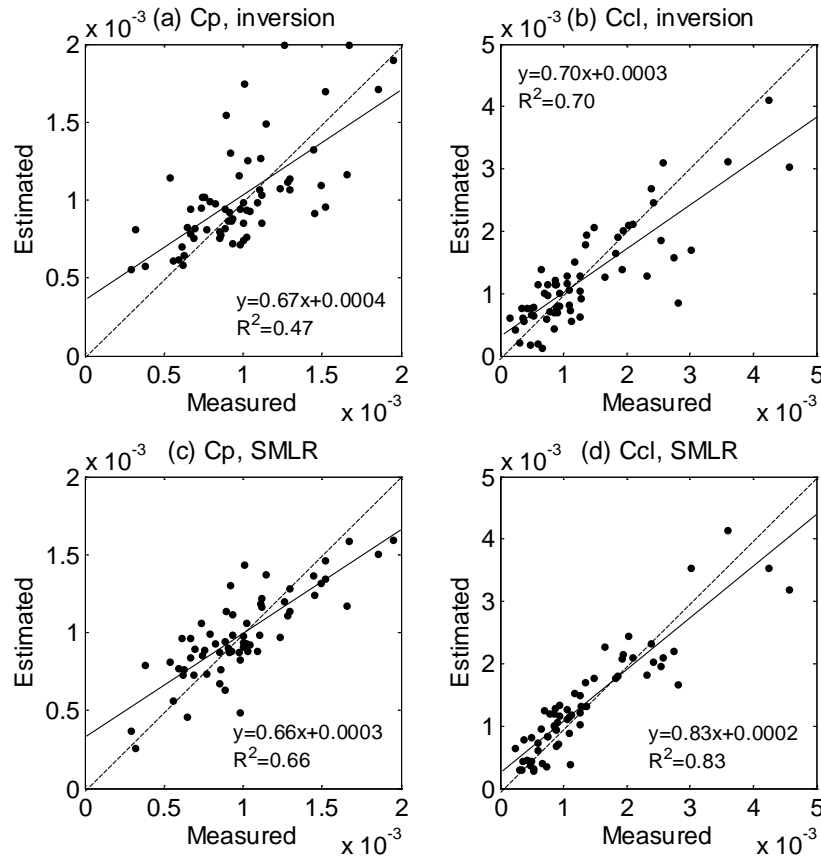
structure parameter and leaf water, bring more uncertainties in the retrieval of protein and cellulose + lignin. Several studies have demonstrated that selecting spectral subsets will return a higher accuracy than using full wavelengths when inverting radiative transfer models (Darvishzadeh et al. 2008; Meroni et al. 2004; Schlerf and Atzberger 2006). The reasons for this can be poorly measured spectra, as some wavelengths may not be well described in the model (Schlerf and Atzberger 2006), or extra bands may add noise without adding significant information (Weiss et al. 2000). We alleviated the “ill-posed” inverse problem and uncertainties in estimation by selecting specific spectral subsets for cellulose + lignin and protein. The selected spectral subsets improved the accuracy of estimating leaf protein and cellulose + lignin when compared to using full wavelengths. The accuracy of retrieved protein and cellulose + lignin from PROSPECT model inversion was comparable with that from stepwise multiple linear regression as calculated in Chapter 3 (Fig. 6.4).



**Figure 6.2** Recalibrated specific absorption coefficients of protein (a) and cellulose + lignin (b) determined on fresh leaves. The vertical dashed lines correspond to published absorption features of protein (a) and cellulose + lignin (b) (see details in Table 3.4 and 3.5).



**Figure 6.3** RMSEP, BIAS and SEPC of the modeling of the reflectance and transmittance of fresh leaves by this study, Jacquemoud et al. (1996) and PROSPECT-5. RMSEP, BIAS and SEPC of this study were averaged values from 100 repetitions of k-fold cross validation.



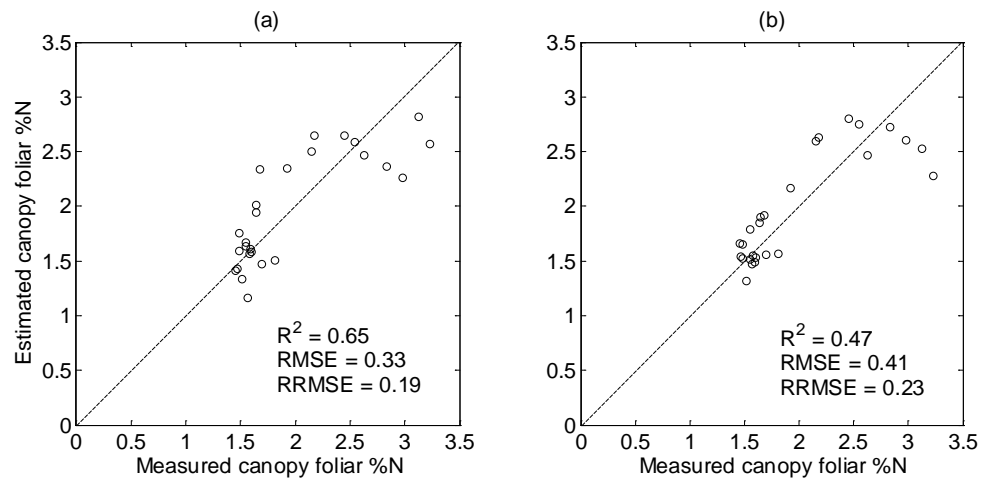
**Figure 6.4** Measured versus estimated protein content (Cp, g/cm<sup>2</sup>) and cellulose + lignin content (Ccl, g/cm<sup>2</sup>) for fresh leaves using PROSPECT-5 inversion with recalibrated specific absorption coefficients and stepwise multiple linear regression (SMLR).

## 6.4 Nitrogen detectable from canopy BRF correcting for canopy structural effects

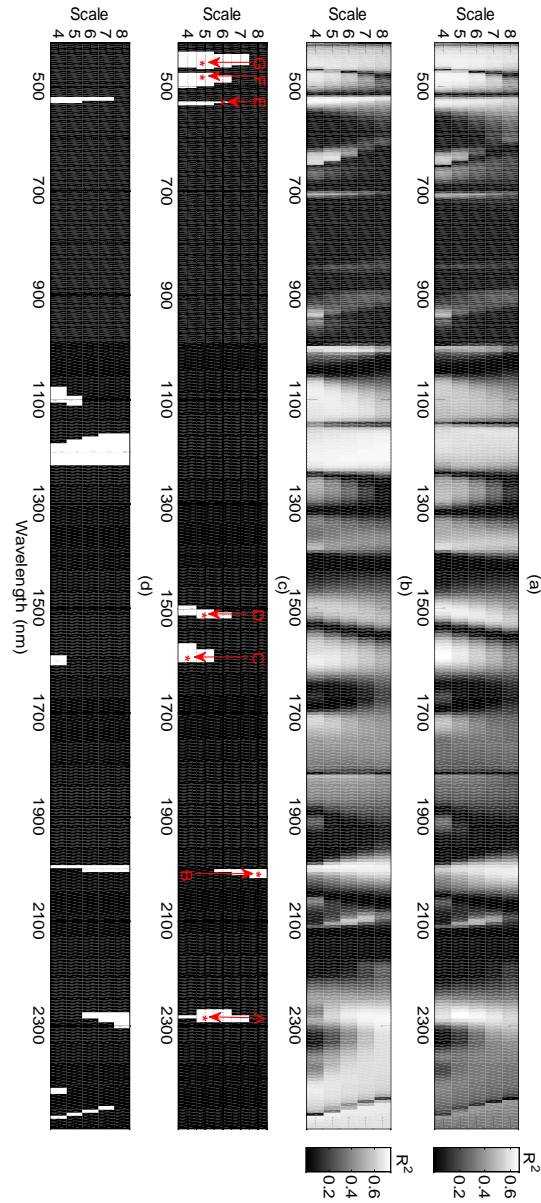
Hyperspectral remote sensing of foliar nitrogen may be confounded by canopy structure which causes more variations in canopy BRF (Knyazikhin et al. 2013a). Estimation of canopy foliar %N has been tested by using the ratio of BRF and DASF spectra (canopy scattering coefficient,  $W_\lambda$ ) in order to suppress the impact of canopy structure (Knyazikhin et al. 2013a), however, their analyses were restricted to using information from each wavelength between 423 and 855 nm. The estimation of %N using the scattering coefficient has not been investigated for longer spectral wavelengths (> 855 nm), where many nitrogen absorption bands are located in (Curran 1989; Fourty et al. 1996; Kokaly et al. 2009).

Chapter 4 derived the canopy scattering coefficient by correcting canopy structure effects on canopy BRF over the full spectrum from 400 to 2500 nm using airborne hyperspectral data, and then applied a continuous wavelet analysis (CWA) to the scattering coefficient in order to estimate %N (Wang et al. 2016b). Results confirmed the feasibility of accurately estimating canopy foliar %N ( $R^2 = 0.65$ , RMSE = 0.33) from the scattering coefficient using a wavelet based approach, and that the wavelet based approach yielded a higher accuracy when compared to a standard approach partial least square regression ( $R^2 = 0.47$ , RMSE = 0.41) (Fig. 6.5). The spectral invariants variable, DASF, is derived based on physical laws and serves as an efficient mechanism for correcting the impact of canopy structure on canopy BRF. Canopy foliar nitrogen can be accurately estimated from the canopy scattering coefficient which suppressed the impact of canopy structure. Additionally, the continuous wavelet analysis was capable of identifying the wavelet features most sensitive to nitrogen variations which can be used for %N estimation. Results showed that the wavelet features most sensitive to %N variation in the visible region relate to chlorophyll absorption, while wavelet features in the shortwave infrared regions relate to protein and dry matter absorption (Fig. 6.6). To our knowledge, this is the first study to evaluate the feasibility of wavelet based approach for estimating canopy foliar %N using remote sensing (airborne hyperspectral data).

After correcting for the canopy structural effects on canopy BRF, the retrieval of nitrogen remains a challenge in terms of the confounding effects of leaf properties. Estimation of nitrogen using hyperspectral remote sensing relies on the fact that nitrogen in leaves cause variation in reflectance. Both direct and indirect effects of nitrogen are exerted on reflectance. Since nitrogen is only a relatively small constituent (0.2-6.4%) of leaves (Wright et al. 2004), the direct effect of nitrogen on reflectance through interacting with photons is minimal (Jacquemoud et al. 1996). In addition, the absorption features of nitrogen overlap with those of dry matter and water. Due to the co-varying relationship of nitrogen and chlorophyll/water/dry matter, the wavelet features most sensitive to nitrogen variation may possibly be due to the absorption of chlorophyll/water/dry matter rather than nitrogen. In this study, the two wavelet features  $WP_{1507,5}$  and  $WP_{2275,5}$  were more related to nitrogen. The indirect effects of nitrogen on reflectance may be attributed to the associations between nitrogen, dry matter, water and chlorophyll. For instance, the other five selected wavelet features were subject to the collective effects of the absorption of these constituents.



**Figure 6.5** Canopy foliar %N estimated from by (a) regression models using a combination of wavelet features  $WP_{2005,8}$ ,  $WP_{1507,5}$ ,  $WP_{528,6}$  and  $WP_{477,5}$  derived from the scattering coefficient, and (b) partial least squares regression model.  $R^2$  and RMSE in the figure are cross-validated values.



**Figure 6.6** Correlations scalograms representing the correlation ( $R^2$ ) between wavelet features and (a) canopy foliar %N and (b) canopy dry matter content. (c) and (d) delineate the top 5% wavelet features correlated with %N and canopy dry matter content, respectively. The seven selected wavelet features (A-G) sensitive to %N in (c) are detailed in Table 6.2. The wavelet power was derived from the canopy scattering coefficient.



**Table 6.2** Performance of regression models for estimating canopy foliar %N using different wavelet features from the canopy scattering coefficient. Related features can be referred to Curran (1989).

Code	Wavelet features	$R^2_{cv}$	RMSE <sub>cv</sub>	NRMSE <sub>cv</sub>	Features related to
A	WP <sub>2275,5</sub>	0.52	0.39	0.22	Protein and nitrogen at 2300 nm
B	WP <sub>2005,8</sub>	0.53	0.39	0.22	Water at 1940 nm
C	WP <sub>1591,4</sub>	0.59	0.36	0.20	Starch and sugar at 1580 nm
D	WP <sub>1507,5</sub>	0.54	0.39	0.22	Protein and nitrogen at 1510 nm
E	WP <sub>528,6</sub>	0.54	0.38	0.22	-
F	WP <sub>477,5</sub>	0.56	0.38	0.21	Chlorophyll b at 460 nm
G	WP <sub>451,5</sub>	0.55	0.38	0.21	Chlorophyll b at 460 nm
	WP <sub>combo</sub> (B, D, E, F)	0.65	0.34	0.19	-

## 6.5 Mapping leaf and canopy nitrogen content by inversion of coupled leaf-canopy radiative transfer models

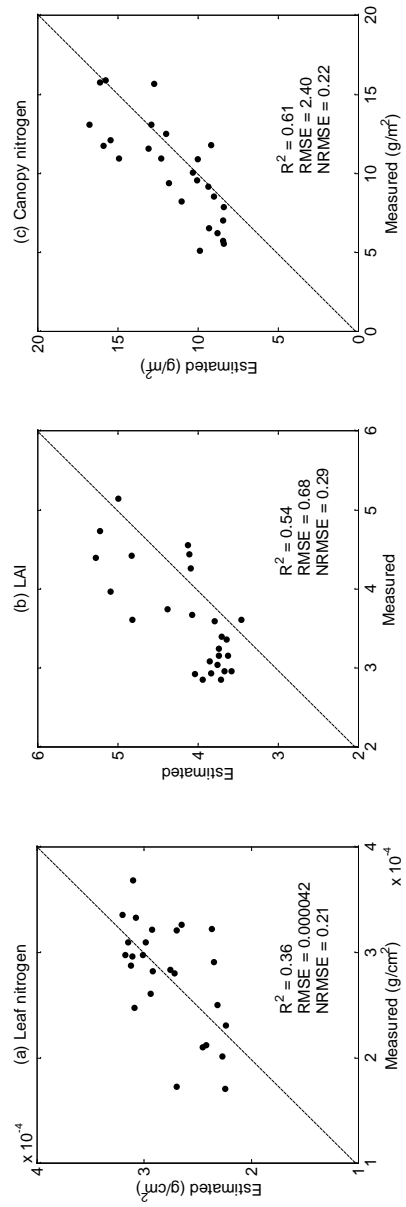
A recalibrated leaf radiative transfer model PROSPECT-5 which incorporated the effects of leaf protein and cellulose + lignin on leaf reflectance and transmittance can be used to estimate nitrogen from fresh leaf spectra via inversion. However, this model has not been tested for the estimation of leaf nitrogen at the canopy level. When estimating leaf nitrogen from the canopy spectra, factors such as canopy structure, illumination/viewing geometry and background further decreased the ability of detecting nitrogen (Asner 1998; Zarco-Tejada et al. 2001).

Chapter 5 evaluated the performance of scaling-up approaches to estimate leaf and canopy nitrogen content from airborne hyperspectral imagery by coupling the revised PROSPECT-5 (Wang et al. 2015a) with a canopy reflectance model. At the canopy level, leaf water and canopy structural parameters such as stem density and crown diameter exerted greater influence on the variability of canopy reflectance than nitrogen, according to a global sensitivity analysis. Despite this, the study confirmed the feasibility of retrieving leaf nitrogen content by scaling up from leaf to canopy level through linking a leaf radiative transfer model PROSPECT-5 and a canopy reflectance model.

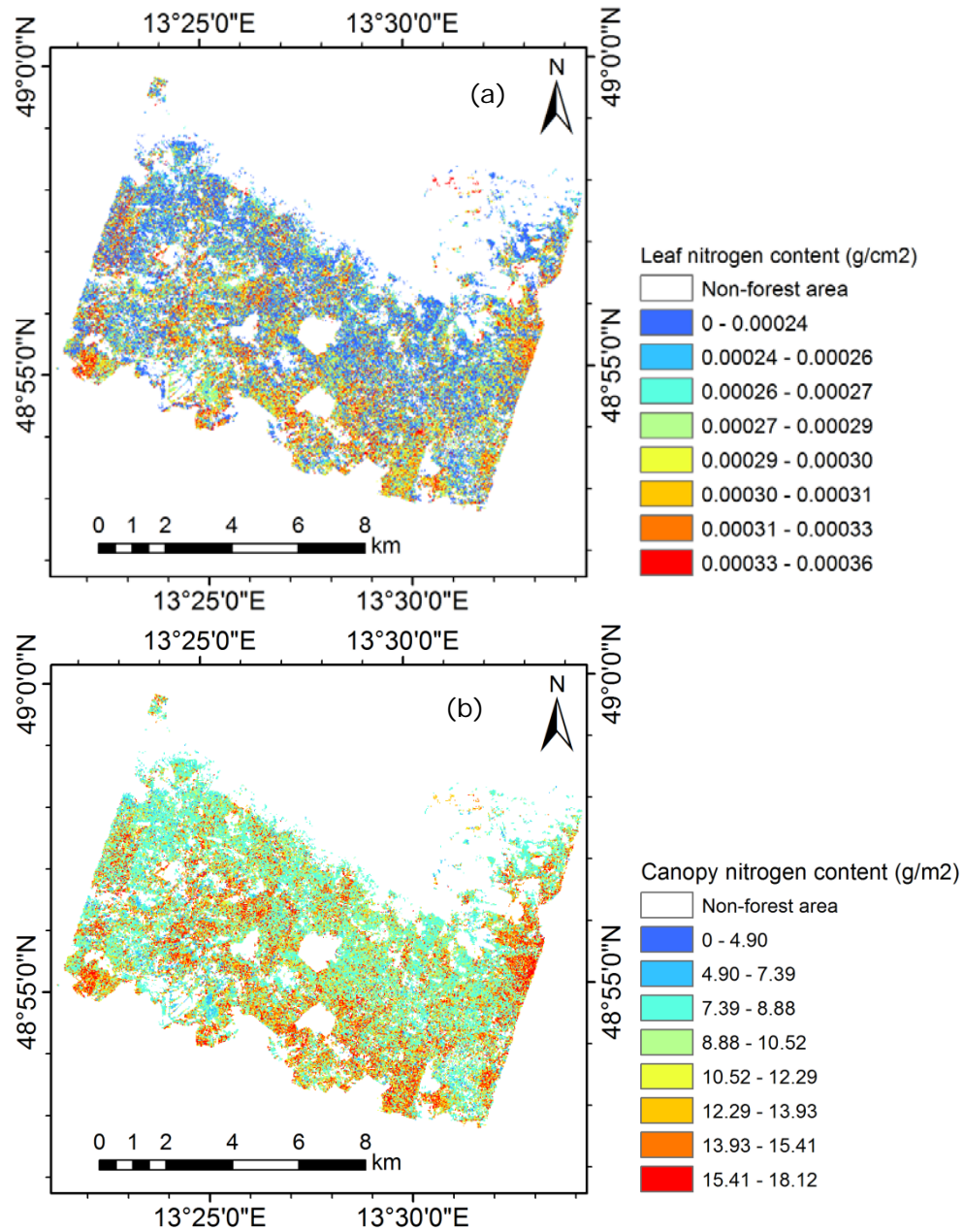
One reason for the feasibility of estimating leaf nitrogen content at the canopy level is the strong multiple scattering effects in forests which can enhance the leaf biochemical signals by up to a factor of two (Baret et al. 1994). In other words,

due to relatively high LAI, the light scattering and absorption enhances the spectroscopic difference among canopies and, hence, leaf biochemistry retrieval (Asner and Martin 2008b). The second explanation is that the parameterization of the INFORM model is probably quite realistic. The INFORM model included parameters for characterizing the canopy geometrical structure, and incorporated the forest light interaction model (FLIM) model to account for the 3-dimensional effects such as shadows, the hotspot and the clumping of leaves in crowns (Schlerf and Atzberger 2006, 2012). These enable the disassociation of the remotely sensed signal from canopy structural effects and leaf optical properties.

More accurate estimation was obtained for canopy nitrogen content ( $R^2 = 0.61$ , RMSE = 2.40, NRMSE = 0.22) than leaf nitrogen content ( $R^2 = 0.38$ , RMSE = 0.000040, NRMSE = 0.20) (Fig. 6.7). The results are consistent with previous studies which have demonstrated that poorer accuracies were obtained for leaf biochemistry compared with canopy biochemistry (Darvishzadeh et al. 2008; Omari et al. 2013; Si et al. 2012). Regulation techniques, such as using prior information, ecological constraints, spectral subsets, and statistical parameters of a certain number of best solutions improved the estimation accuracy. The maps of leaf and canopy nitrogen content (Fig. 6.8) are first to be generated using inversion of coupled leaf-canopy models, and the spatial variation of nitrogen appears to be reasonable and consistent with ecological knowledge.



**Figure 6.7** Measured versus estimated parameters: (a) leaf nitrogen content (b) LAI and (c) canopy nitrogen content from measured spectra using spectral subset 2 by INFORM inversion.



**Figure 6.8** Leaf nitrogen content (a) and canopy nitrogen content (b) mapped using the HySpex airborne image data and inversion of the INFORM in Bavarian Forest National Park, Germany.

## 6.6 Convergence of leaf properties and canopy properties for foliar nitrogen retrieval

Canopy structure confounds the estimation of foliar nitrogen when using canopy spectral data because it is the main driver of canopy reflectance variations. Ollinger et al. (2008) reported that the significant correlation between NIR reflectance (800 – 850 nm) and canopy foliar mass-based nitrogen concentration (%N) can be used for predicting nitrogen. However, Knyazikhin et al. (2013a) pointed out that the relationship can be attributed to the correlation between NIR reflectance and canopy structure. Ollinger et al. (2013) argued that their hypotheses were based on the biological associations between nitrogen and structural traits which affect NIR scattering and reflectance. We claim that canopy structure should not be considered as a problem for nitrogen estimation but rather represents an opportunity for improved accuracy in nitrogen estimation.

First of all, the %N-NIR relationship is not necessarily spurious (Townsend et al. (2013)), because the canopy structure and leaf properties may co-vary across plant functional types (Wright et al. (2004) and Ollinger (2011)). Species and plant functional types (*i.e.* broadleaf and coniferous forest types) account for most of variance of canopy chemistry which has been demonstrated across tropical (Asner et al. 2015; Asner et al. 2009; Asner et al. 2012; Asner et al. 2014), temperate (McNeil et al. 2008) and boreal forests (Gökkaya et al. 2015a; Gökkaya et al. 2015b) as well as Mediterranean ecosystems (Dahlin et al. 2013). The link between species and canopy biochemistry can be explained by the concept of ‘global leaf economics spectrum’ (Wright et al. 2004), which means that the key plant traits such as leaf mass per area, specific leaf area, leaf nitrogen, leaf phosphorous, leaf lifespan, and photosynthesis fall into a spectrum across plant species, and species converge towards the functional traits globally (Reich et al. 1997; Wright et al. 2004).

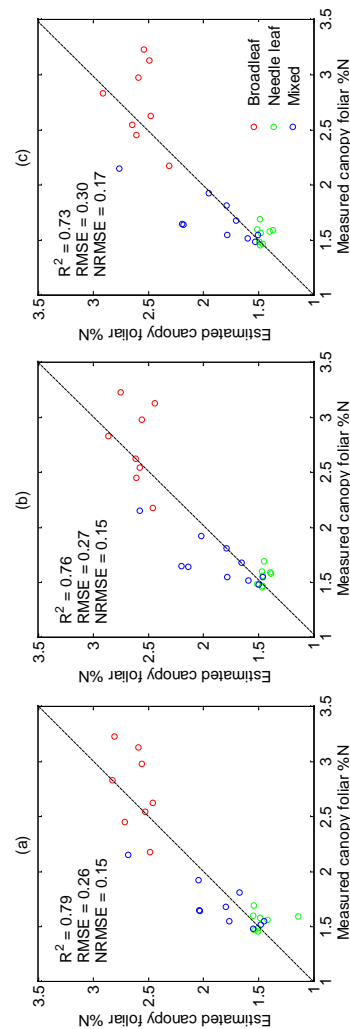
The covariance of these functional traits was evaluated for nitrogen estimation in a mixed temperate forest using three categories of vegetation indices related to biochemical and physical properties of vegetation (*i.e.* nitrogen, structure and chlorophyll) derived from airborne hyperspectral imagery (Wang et al. 2016a). In this mixed temperate forest, functional type and species composition played a dominant role in explaining the variance of canopy foliar nitrogen. This is consistent with findings from different ecosystems, such as temperate, tropical, boreal and Mediterranean ecosystems (Asner et al. 2009; Asner et al. 2012; Asner et al. 2014; Dahlin et al. 2013; Gökkaya et al. 2015a; McNeil et al. 2008).

Comparably accurate estimations of %N were observed across all three categories of vegetation indices (Fig. 6.9). The best performing nitrogen related indices utilized the physical basis of nitrogen absorption features in canopy reflectance. The structural indices, e.g. the mean NIR reflectance between 800 and 850 nm provided an accurate estimation of nitrogen, which was in agreement with the findings reported for some temperate and boreal forests in North America (Ollinger et al. 2008). In addition to NIR reflectance, most of the structural related indices were capable of estimating leaf nitrogen at a moderate to good accuracy. The results might be explained by the control of functional type and species composition with regard to the ‘global leaf economics spectra’ types (Wright et al. (2004)). That is the functional convergence across species among optically important leaf traits such as leaf mass per area, nitrogen concentration, and canopy structural properties such as LAI (Ollinger 2011). The best performing chlorophyll related indices, used the red-edge region and were subjected to combined influences of the strong chlorophyll absorption and structural properties.

Secondly, Chapter 3 demonstrated that nitrogen can still be detected while using canopy BRF to correct for canopy structural effects, though with lower accuracies. It has been argued that nitrogen influences canopy reflectance in both direct and indirect ways (Lepine et al. 2016). The direct influence is from the absorption features of nitrogen-contained constituents such as proteins and chlorophylls. This signal is weak when detected at the canopy level due to the water absorption mask as well as multiple scattering processes due to canopy structural effects. The indirect influence of nitrogen on canopy reflectance is exerted due to the association between nitrogen and canopy structure. Previous studies confirmed this by showing the importance of the near-infrared spectral regions in nitrogen estimation using partial least squares regression (Gökkaya et al. 2015a; Lepine et al. 2016; Martin et al. 2008). Similar conclusions can be drawn from a recent study using vegetation indices (Wang et al. 2016a). For the sake of accurate nitrogen estimates, the association between nitrogen and canopy structure is recommended to be used for retrieving nitrogen in a mixed temperate and boreal forest.

Thirdly, mapping forest canopy foliar nitrogen concentration (%N) at a local level is practical with the aid of imaging spectrometry. The current lack of higher spatial resolution satellite hyperspectral sensors impedes the mapping of nitrogen at larger scales (e.g. continental or global scale), though the launch of EnMAP mission planned for 2018 may provide an opportunity. Ollinger et al. (2008)

provided a continental-scale map of foliar nitrogen by utilizing a statistical relationship between %N and canopy bidirectional reflectance factor (BRF) for near-infrared (NIR) wavelengths (800 - 850 nm) for temperate and boreal forests. The findings were encouraging for the remote sensing community because they allow the estimation of leaf nitrogen across larger extents and potentially frequent updates using data from broadband satellite data such as MODIS (Ollinger et al. 2008). Therefore, the functional link between %N and NIR reflectance could be utilized as a simple and rapid means of generating regional to continual maps of nitrogen.



**Figure 6.9** The 1:1 relationship between the estimated and measured canopy foliar %N based on best performing vegetation indices (a) NDNI<sub>1510</sub> (b) Boochs2 and (c) NIR BRF, 800-850 nm.  $R^2$ , RMSE and NRMSE are values from leave-one-out cross validation.

## **6.7 Broader implications of nitrogen estimation for ecological studies**

This thesis demonstrates an approach for estimating foliar nitrogen at the canopy level by coupling leaf and canopy radiative transfer models. The robustness and transferability of radiative transfer models suggests that the approaches might be transferred to other sites. The combination of spectral invariants theory and continuous wavelet analysis have been shown to be an efficient and robust way of hyperspectral remote sensing of canopy foliar nitrogen regardless of the confounding effects of canopy structure. The approaches are applicable to estimation of all leaf biochemistry, such as leaf water, dry matter, cellulose, lignin, and sun-induced chlorophyll fluorescence (Damm et al. 2015; Knyazikhin et al. 2013a), since the decoupling of canopy structure in canopy BRF is a common challenge. The covariance of leaf nitrogen, leaf traits, canopy structure parameters were also investigated and recommended for nitrogen estimation.

Accurate estimation of foliar nitrogen is beneficial in a variety of ecological applications. Global maps of foliar nitrogen content could be generated when hyperspectral satellites such as Sentinel (Malenovsky et al. 2012), EnMAP (Kaufmann et al. 2008) and HypSPIRI (<http://hyspiri.jpl.nasa.gov/>) become operational. In turn, such monitoring would support the development and demonstration of Essential Biodiversity Variables (EBVS), which are used for assessing biodiversity (Pereira et al. 2013; Skidmore et al. 2015). Regional and global mapping of leaf nitrogen will further our understanding of the photosynthesis process, net primary productivity, and carbon dynamics (Evans 1989; Heimann and Reichstein 2008; Ollinger and Smith 2005; Reich et al. 2006).

Canopy level protein (equivalent to nitrogen) (Jacquemoud et al. 1996) and cellulose + lignin (approximation of carbon) would not only yield valuable information about photosynthesis and net production, but would also allow the C:N ratio to be estimated, which can be used for evaluating the litter decomposition and nutrient cycling (Aber and Federer 1992; Steudler et al. 1989). In a broader context, foliar nitrogen as an input to ecosystem process models is essential in understanding the nutrient cycle and climatic feedbacks (Singh et al. 2015; Zaehle et al. 2014). Moreover, due to the elevated anthropogenic nitrogen sources (Hogberg 2007; Vitousek et al. 1997), it is critical to examine the foliar nitrogen responses in forest ecosystems to atmospheric nitrogen deposition (McNeil et al. 2007). Such responses may allow for monitoring dynamics of



forest carbon sequestration, nitrogen cycling, species composition and biodiversity, and may provide valuable information for air pollution control.

## 6.8 Future research

Future research can be undertaken along two lines, i.e. further development of radiative transfer models and ecological understanding.

From a RTM model perspective, more accurate estimation may benefit from recalibrated leaf radiative transfer model PROSPECT specifically for conifer needle species if there are available datasets. Besides, the recalibrated PROSPECT-5 in this thesis considered the wavelengths between 800 – 2500 nm, further research needs to be conducted to incorporate the special interval of 400 – 799 nm by considering the chlorophyll-nitrogen relationship. Empirical research has showed the importance of the red-edge information in nitrogen estimation, thus should also be involved in the radiative transfer models. Different spectral transformation methods have been used to reduce the influence of water on estimating leaf biochemical constituents in statistical studies, such as first or second derivatives, continuum removal (Curran et al. 2001; Huang et al. 2004; Kokaly and Clark 1999; Mutanga and Skidmore 2004; Mutanga et al. 2005), and water removal through spectral matching (Gao and Goetz 1994; Ramoelo et al. 2011; Schlerf et al. 2010). It is worthwhile to investigate such techniques specifically for inversion of a radiative transfer model.

To alleviate the “ill-posed” inverse problem, inversion approaches such as using prior information (Combal et al. 2002), spectral sampling sets (Darvishzadeh et al. 2008; Wang et al. 2015a; Weiss et al. 2000), and “step-by-step” inversion (Li and Wang 2011) were involved in this thesis. More inversion techniques need to be tested, such as using temporal and spatial information (Houborg et al. 2009; Houborg et al. 2015) as constraints. With the aid of light detection and ranging (Lidar), canopy structural parameters may be input to canopy radiative transfer models and used as constraints during model inversion (Asner et al. 2015; Combal et al. 2002; Gökkaya et al. 2015a; Niinemets 2010). Moreover, empirical relationships can complement radiative transfer models, which may in turn facilitate the development of improved radiative transfer models. Also, empirical relationships can be utilized as constraints during both the forward modeling and the inversion process.

From an ecological perspective, a deeper understanding of the convergence between nitrogen, leaf traits and canopy structural parameters is needed. Research

could be conducted for more species and ecosystems, in more heterogeneous forest areas as well as across larger extents. Furthermore, the physical interpretations behind the correlation between nitrogen and hyperspectral data could be explored.

## Bibliography

- Aber, J., & Federer, C.A. (1992). A generalized, lumped-parameter model of photosynthesis, evapotranspiration and net primary production in temperate and boreal forest ecosystems. *Oecologia*, 92, 463-474
- Ali, A.M., Skidmore, A.K., Darvishzadeh, R., Duren, I.v., Holzwarth, S., & Müller, J. (2016a). Retrieval of forest leaf functional traits from HySpex imagery using radiative transfer models and continuous wavelet analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, 122, 68-80.
- Ali, A.M., Darvishzadeh, R., Skidmore, A.K., & Duren, I.v. (2016b). Effects of canopy structural variables on retrieval of Leaf dry matter content and specific leaf area from remotely sensed data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9, 898-909
- Ali, A.M., Darvishzadeh, R., Skidmore, A.K., Duren, I.v., Heiden, U., & Heurich, M. (2016c). Estimating leaf functional traits by inversion of PROSPECT: Assessing leaf dry matter content and specific leaf area in mixed mountainous forest. *International Journal of Applied Earth Observation and Geoinformation*, 45, Part A, 66-76
- AOAC (1970). *Official Methods of Analysis*. (11th ed.). Washington, DC: Association of Official Analytical Chemists
- Asbjornsen, H., Goldsmith, G.R., Alvarado-Barrientos, M.S., Rebel, K., Van Osch, F.P., Rietkerk, M., Chen, J., Gotsch, S., Tobón, C., Geissert, D.R., Gómez-Tagle, A., Vache, K., & Dawson, T.E. (2011). Ecohydrological advances and applications in plant–water relations research: a review. *Journal of Plant Ecology*, 4, 3-22
- ASD (2008). Integrating Sphere User Manual. In, *ASD.Document 600660 Rev. A*. Analytical Spectral Devices, Inc. (ASD), Boulder, CO, USA
- Asner, G., Martin, R., Carlson, K., Rascher, U., & Vitousek, P. (2006). Vegetation–Climate Interactions among Native and Invasive Species in Hawaiian Rainforest. *Ecosystems*, 9, 1106-1117
- Asner, G.P. (1998). Biophysical and Biochemical Sources of Variability in Canopy Reflectance. *Remote Sensing of Environment*, 64, 234-253
- Asner, G.P., & Martin, R.E. (2008a). Airborne spectranomics: mapping canopy chemical and taxonomic diversity in tropical forests. *Frontiers in Ecology and the Environment*, 7, 269-276
- Asner, G.P., & Martin, R.E. (2008b). Spectral and chemical analysis of tropical forests: Scaling from leaf to canopy levels. *Remote Sensing of Environment*, 112, 3958-3970

- Asner, G.P., Martin, R.E., Anderson, C.B., & Knapp, D.E. (2015). Quantifying forest canopy traits: Imaging spectroscopy versus field survey. *Remote Sensing of Environment*, 158, 15-27
- Asner, G.P., Martin, R.E., Ford, A.J., Metcalfe, D.J., & Liddell, M.J. (2009). Leaf chemical and spectral diversity in Australian tropical forests. *Ecological Applications*, 19, 236-253
- Asner, G.P., Martin, R.E., Knapp, D.E., Tupayachi, R., Anderson, C., Carranza, L., Martinez, P., Houcheime, M., Sinca, F., & Weiss, P. (2011). Spectroscopy of canopy chemicals in humid tropical forests. *Remote Sensing of Environment*, 115, 3587-3598
- Asner, G.P., Martin, R.E., & Suhaili, A.B. (2012). Sources of canopy chemical and spectral diversity in lowland Bornean forest. *Ecosystems*, 15, 504-517
- Asner, G.P., Martin, R.E., Tupayachi, R., Anderson, C.B., Sinca, F., Carranza-Jiménez, L., & Martinez, P. (2014). Amazonian functional diversity from forest canopy chemical assembly. *Proceedings of the National Academy of Sciences of the United States of America*
- Asner, G.P., & Vitousek, P.M. (2005). Remote analysis of biological invasion and biogeochemical change. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 4383-4386
- Atzberger, C. (2000). *Development of an invertible forest reflectance model: The INFOR-model*.
- Atzberger, C., Jarmer, T., Schlerf, M., Kötz, B., & Werner, W. (2003). Spectroradiometric determination of wheat bio-physical variables: comparison of different empirical-statistical approaches. In R. Goossens (Ed.), *Remote Sensing in Transitions, Proc. 23rd EARSeL symposium, Belgium* (pp. 463-470). Belgium
- Axelsson, C., Skidmore, A.K., Schlerf, M., Fauzi, A., & Verhoef, W. (2013). Hyperspectral analysis of mangrove foliar chemistry using PLSR and support vector regression. *International Journal of Remote Sensing*, 34, 1724-1743
- Baldocchi, D.D., Xu, L., & Kiang, N. (2004). How plant functional-type, weather, seasonal drought, and soil physical properties alter water and energy fluxes of an oak-grass savanna and an annual grassland. *Agricultural and Forest Meteorology*, 123, 13-39
- Banskota, A., Wynne, R.H., Thomas, V.A., Serbin, S.P., Kayastha, N., Gastellu-Etchegorry, J.P., & Townsend, P.A. (2013). Investigating the utility of wavelet transforms for inverting a 3-D radiative transfer model using hyperspectral data to retrieve forest LAI. *Remote Sensing*, 5, 2639-2659

- Baret, F., & Fourty, T. (1997a). Estimation of leaf water content and specific leaf weight from reflectance and transmittance measurements. *Agronomie*, 17, 455-464
- Baret, F., & Fourty, T.H. (1997b). Radiometric Estimates of Nitrogen Status of Leaves and Canopies. In G. Lemaire (Ed.), *Diagnosis of the Nitrogen Status in Crops* (pp. 201-227): Springer Berlin Heidelberg
- Baret, F., Houlès, V., & Guérif, M. (2007). Quantification of plant stress using remote sensing observations and crop models: the case of nitrogen management. *Journal of Experimental Botany*, 58, 869-880
- Baret, F., Vanderbilt, V.C., Steven, M.D., & Jacquemoud, S. (1994). Use of spectral analogy to evaluate canopy reflectance sensitivity to leaf optical properties. *Remote Sensing of Environment*, 48, 253-260
- Barton, F.E. (1987). *Analytical application to fibrous foods and commodities*. St. Paul, MN: American Association of Cereal Chemists
- Bian, M. (2013). Assessing the quality of tea by hyperspectral techniques. In (p. 143). Enschede, the Netherlands: University of Twente Faculty of Geo-Information and Earth Observation (ITC)
- Blackburn, G.A., & Ferwerda, J.G. (2008). Retrieval of chlorophyll concentration from leaf reflectance spectra using wavelet analysis. *Remote Sensing of Environment*, 112, 1614-1632
- Botha, E.J., Zebarth, B.J., & Leblon, B. (2006). Non-destructive estimation of potato leaf chlorophyll and protein contents from hyperspectral measurements using the PROSPECT radiative transfer model. *Canadian Journal of Plant Science*, 86, 279-291
- Bowyer, P., & Danson, F.M. (2004). Sensitivity of spectral reflectance to variation in live fuel moisture content at leaf and canopy level. *Remote Sensing of Environment*, 92, 297-308
- Bring, J. (1994). How to Standardize Regression Coefficients. *The American Statistician*, 48, 209-213
- Broge, N.H., & Leblanc, E. (2001). Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sensing of Environment*, 76, 156-172
- Bruce, L.M., Koger, C.H., & Jiang, L. (2002). Dimensionality reduction of hyperspectral data using discrete wavelet transform feature extraction. *IEEE Transactions on Geoscience and Remote Sensing*, 40, 2331-2338

- Bruce, L.M., Morgan, C., & Larsen, S. (2001). Automated detection of subpixel hyperspectral targets with continuous and discrete wavelet transforms. *IEEE Transactions on Geoscience and Remote Sensing*, 39, 2217-2226
- Buiteveld, H., Hakvoort, J.H.M., & Donze, M. (1994). The optical properties of pure water. *Ocean Optics XIII*, 2258, 174-183
- Cannavó, F. (2012). Sensitivity analysis for volcanic source modeling quality assessment and model selection. *Computers & Geosciences*, 44, 52-59
- Card, D.H., Peterson, D.L., Matson, P.A., & Aber, J.D. (1988). Prediction of leaf chemistry by the use of visible and near infrared reflectance spectroscopy. *Remote Sensing of Environment*, 26, 123-147
- Ceccato, P., Flasse, S., & Grégoire, J.M. (2002). Designing a spectral index to estimate vegetation water content from remote sensing data: Part 1: Theoretical approach. *Remote Sensing of Environment*, 82, 188-197
- Chapin, F.S., III, Bloom, A.J., Field, C.B., & Waring, R.H. (1987). Plant responses to multiple environmental factors. *Bioscience*, 37, 49-57
- Chen, P.F., Wang, J.H., Huang, W.J., Tremblay, N., Ou, Y.Z., & Zhang, Q. (2013). Critical Nitrogen Curve and Remote Detection of Nitrogen Nutrition Index for Corn in the Northwestern Plain of Shandong Province, China. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6, 682-689
- Cheng, T., Rivard, B., Sanchez-Azofeifa, A.G., Feret, J.B., Jacquemoud, S., & Ustin, S.L. (2014a). Deriving leaf mass per area (LMA) from foliar reflectance across a variety of plant species using continuous wavelet analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, 87, 28-38
- Cheng, T., Riaño, D., & Ustin, S.L. (2014b). Detecting diurnal and seasonal variation in canopy water content of nut tree orchards from airborne imaging spectroscopy data using continuous wavelet analysis. *Remote Sensing of Environment*, 143, 39-53
- Cheng, T., Rivard, B., & Sanchez-Azofeifa, A. (2011). Spectroscopic determination of leaf water content using continuous wavelet analysis. *Remote Sensing of Environment*, 115, 659-670
- Cho, M. (2007a). *Hyperspectral remote sensing of biochemical and biophysical parameters: the derivate red-edge" double-peak feature", a nuisance or an opportunity? : [SI: sn]*
- Cho, M.A. (2007b). Hyperspectral remote sensing of biochemical and biophysical parameters: the derivate red-edge" double-peak feature", a nuisance or an opportunity? In (p. 206). Enschede, the Netherlands:

- University of Twente Faculty of Geo-Information and Earth Observation (ITC)
- Clark, R.N., & Roush, T.L. (1984). Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications. *Journal of Geophysical Research: Solid Earth*, 89, 6329-6340
- Clevers, J., Kooistra, L., & Schalepman, M.E. (2008). Using spectral information from the NIR water absorption features for the retrieval of canopy water content. *International Journal of Applied Earth Observation and Geoinformation*, 10, 388-397
- Clevers, J.G.P.W., & Gitelson, A.A. (2013). Remote estimation of crop and grass chlorophyll and nitrogen content using red-edge bands on Sentinel-2 and -3. *International Journal of Applied Earth Observation and Geoinformation*, 23, 344-351
- Clevers, J.G.P.W., & Kooistra, L. (2012). Using hyperspectral remote sensing data for retrieving canopy chlorophyll and nitrogen content. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5, 574-583
- Collatz, G., Ribas-Carbo, M., & Berry, J. (1992). Coupled Photosynthesis-Stomatal Conductance Model for Leaves of C4 Plants. *Functional Plant Biology*, 19, 519-538
- Colombo, R., Merom, M., Marchesi, A., Busetto, L., Rossini, M., Giardino, C., & Panigada, C. (2008). Estimation of leaf and canopy water content in poplar plantations by means of hyperspectral indices and inverse modeling. *Remote Sensing of Environment*, 112, 1820-1834
- Combal, B., Baret, F., Weiss, M., Trubuil, A., Mace, D., Pragnere, A., Myneni, R., Knyazikhin, Y., & Wang, L. (2002). Retrieval of canopy biophysical variables from bidirectional reflectance - Using prior information to solve the ill-posed inverse problem. *Remote Sensing of Environment*, 84, 1-15
- Coops, N.C., Smith, M.L., Martin, M.E., & Ollinger, S.V. (2003). Prediction of eucalypt foliage nitrogen content from satellite-derived hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 41, 1338-1346
- Curran, P.J. (1989). Remote sensing of foliar chemistry. *Remote Sensing of Environment*, 30, 271-278
- Curran, P.J. (2001). Imaging spectrometry for ecological applications. *International Journal of Applied Earth Observation and Geoinformation*, 3, 305-312
- Curran, P.J., Dungan, J.L., & Peterson, D.L. (2001). Estimating the foliar biochemical concentration of leaves with reflectance spectrometry testing

- the Kokaly and Clark methodologies. *Remote Sensing of Environment*, 76, 349-359
- D'Errico, J. (2012). Fminsearchbnd, fminsearchcon. In. <http://www.mathworks.com/matlabcentral/fileexchange/8277-fminsearchbnd--fminsearchcon/content/FMINSEARCHBND/fminsearchbnd.m>
- Dahlin, K.M., Asner, G.P., & Field, C.B. (2013). Environmental and community controls on plant canopy chemistry in a Mediterranean-type ecosystem. *Proceedings of the National Academy of Sciences of the United States of America*, 110, 6895-6900
- Damm, A., Guanter, L., Paul-Limoges, E., van der Tol, C., Hueni, A., Buchmann, N., Eugster, W., Ammann, C., & Schaepman, M.E. (2015). Far-red sun-induced chlorophyll fluorescence shows ecosystem-specific relationships to gross primary production: An assessment based on observational and modeling approaches. *Remote Sensing of Environment*, 166, 91-105
- Danks, S.M., Evans, E.H., & Whittaker, P.A. (1983). *Photosynthetic Systems: Structure, Function and Assembly* (p. 37). New York: Wiley
- Darvishzadeh, R., Atzberger, C., Skidmore, A., & Schlerf, M. (2011). Mapping grassland leaf area index with airborne hyperspectral imagery: A comparison study of statistical approaches and inversion of radiative transfer models. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66, 894-906
- Darvishzadeh, R., Skidmore, A., Schlerf, M., & Atzberger, C. (2008). Inversion of a radiative transfer model for estimating vegetation LAI and chlorophyll in a heterogeneous grassland. *Remote Sensing of Environment*, 112, 2592-2604
- Dawson, T.P., Curran, P.J., & Plummer, S.E. (1998). LIBERTY - Modeling the effects of Leaf Biochemical Concentration on Reflectance Spectra. *Remote Sensing of Environment*, 65, 50-60
- DeJonge, K.C., Ascough, J.C., Ahmadi, M., Andales, A.A., & Arabi, M. (2012). Global sensitivity and uncertainty analysis of a dynamic agroecosystem model under different irrigation treatments. *Ecological Modelling*, 231, 113-125
- Du, P., Kibbe, W.A., & Lin, S.M. (2006). Improved peak detection in mass spectrum by incorporating continuous wavelet transform-based pattern matching. *Bioinformatics*, 22, 2059-2065
- Elisseeff, A., & Pontil, M. (2002). Leave-one-out Error and Stability of Learning Algorithms with Applications. In J. Suykens, G. Horvath, S. Basu, C.



- Micchelli, & J. Vandewalle (Eds.), *Learning Theory and Practice*: IOS Press
- Ellsworth, D.S., & Reich, P.B. (1993). Canopy structure and vertical patterns of photosynthesis and related leaf traits in a deciduous forest. *Oecologia*, 96, 169-178
- Elvidge, C.D. (1990). Visible and near infrared reflectance characteristics of dry plant materials. *International Journal of Remote Sensing*, 11, 1775-1795
- Evans, J. (1989). Photosynthesis and nitrogen relationships in leaves of C3 plants. *Oecologia*, 78, 9-19
- Evans, J.R. (1983). Nitrogen and photosynthesis in the flag leaf of wheat (*Triticum aestivum* L.). *Plant Physiology*, 72, 297-302
- Evans, J.R., & Poorter, H. (2001). Photosynthetic acclimation of plants to growth irradiance: the relative importance of specific leaf area and nitrogen partitioning in maximizing carbon gain. *Plant, Cell & Environment*, 24, 755-767
- Ewerts (1985). 3ème Directive avec rectification (J.O.C.E. du 27 novembre 1980) In, *Recueil des Normes Françaises-AFNOR*
- Feret, J.-B., Francois, C., Asner, G.P., Gitelson, A.A., Martin, R.E., Bidet, L.P.R., Ustin, S.L., le Maire, G., & Jacquemoud, S. (2008). PROSPECT-4 and 5: Advances in the leaf optical properties model separating photosynthetic pigments. *Remote Sensing of Environment*, 112, 3030-3043
- Ferwerda, J.G., & Jones, S.D. (2006). Continuous wavelet transformations for hyperspectral feature detection. In A. Riedl, W. Kainz, & G.A. Elmes (Eds.), *Progress in Spatial Data Handling* (pp. 167-178). Springer Berlin Heidelberg
- Field, C., & Mooney, H.A. (1986). The photosynthesis-nitrogen relationship in wild plants. In T.J. Givnish (Ed.), *On the Economy of Plant Form and Function* (pp. 25-55). Cambridge: Cambridge University Press
- Foley, S., Rivard, B., Sanchez-Azofeifa, G.A., & Calvo, J. (2006). Foliar spectral properties following leaf clipping and implications for handling techniques. *Remote Sensing of Environment*, 103, 265-275
- Fourty, T., & Baret, F. (1998). On spectral estimates of fresh leaf biochemistry. *International Journal of Remote Sensing*, 19, 1283-1297
- Fourty, T., Baret, F., Jacquemoud, S., Schmuck, G., & Verdebout, J. (1996). Leaf optical properties with explicit description of its biochemical composition: Direct and inverse problems. *Remote Sensing of Environment*, 56, 104-117

- Gamon, J.A., Peñuelas, J., & Field, C.B. (1992). A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sensing of Environment*, 41, 35-44
- Ganapol, B.D., Johnson, L.F., Hammer, P.D., Hlavka, C.A., & Peterson, D.L. (1998). LEAFMOD: A new within-leaf radiative transfer model. *Remote Sensing of Environment*, 63, 182-193
- Gao, B.-C., & Goetz, A.F.H. (1994). Extraction of dry leaf spectral features from reflectance spectra of green vegetation. *Remote Sensing of Environment*, 47, 369-374
- Garnier, E., & Laurent, G. (1994). Leaf anatomy, specific mass and water content in congeneric annual and perennial grass species. *New Phytologist*, 128, 725-736
- Gastellu-Etchegorry, J.P., Demarez, V., Pinel, V., & Zagolski, F. (1996). Modeling radiative transfer in heterogeneous 3-D vegetation canopies. *Remote Sensing of Environment*, 58, 131-156
- Geladi, P., & Kowalski, B.R. (1986). Partial least-squares regression: a tutorial. *Analytica Chimica Acta*, 185, 1-17
- Gökkaya, K., Thomas, V., Noland, T.L., McCaughey, H., Morrison, I., & Treitz, P. (2015a). Prediction of macronutrients at the canopy level using spaceborne imaging spectroscopy and LiDAR data in a mixedwood boreal forest. *Remote Sensing*, 7, 9045-9069
- Gökkaya, K., Thomas, V., Noland, T., McCaughey, H., Morrison, I., & Treitz, P. (2015b). Mapping continuous forest type variation by means of correlating remotely sensed metrics to canopy N:P ratio in a boreal mixedwood forest. *Applied Vegetation Science*, 18, 143-157
- Gower, S.T., Reich, P.B., & Son, Y. (1993). Canopy dynamics and aboveground production of five tree species with different leaf longevities. *Tree Physiology*, 12, 327-345
- Green, D.S., Erickson, J.E., & Kruger, E.L. (2003). Foliar morphology and canopy nitrogen as predictors of light-use efficiency in terrestrial vegetation. *Agricultural and Forest Meteorology*, 115, 163-171
- Guanter, L., Kaufmann, H., Segl, K., Foerster, S., Rogass, C., Chabrillat, S., Kuester, T., Hollstein, A., Rossner, G., & Chlebek, C. (2015). The EnMAP spaceborne imaging spectroscopy mission for earth observation. *Remote Sensing*, 7, 8830-8857
- Hansen, P.M., & Schjoerring, J.K. (2003). Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference

- vegetation indices and partial least squares regression. *Remote Sensing of Environment*, 86, 542-553
- Harrison, M.T., Edwards, E.J., Farquhar, G.D., Nicotra, A.B., & Evans, J.R. (2009). Nitrogen in cell walls of sclerophyllous leaves accounts for little of the variation in photosynthetic nitrogen-use efficiency. *Plant, Cell & Environment*, 32, 259-270
- Heimann, M., & Reichstein, M. (2008). Terrestrial ecosystem carbon dynamics and climate feedbacks. *Nature*, 451, 289-292
- Hergert, H.L. (1971). Infrared spectra. *Lignins: Occurrence, Formation Structure and Reactions*. New York: Wiley-Interscience
- Hernandez-Clemente, R., Navarro-Cerrillo, R.M., & Zarco-Tejada, P.J. (2014). Deriving predictive relationships of carotenoid content at the canopy level in a conifer forest using hyperspectral imagery and model simulation. *IEEE Transactions on Geoscience and Remote Sensing*, 52, 5206-5217
- Heurich, M., Beudert, B., Rall, H., & Křenová, Z. (2010). National parks as model regions for interdisciplinary long-term ecological research: The Bavarian Forest and Šumavá national parks underway to transboundary ecosystem research. In F. Müller, C. Baessler, H. Schubert, & S. Klotz (Eds.), *Long-Term Ecological Research* (pp. 327-344): Springer Science & Business Media
- Heurich, M., & Neufanger, M. (2005). *Die Wälder des Nationalparks Bayerischer Wald. Ergebnisse der Waldinventur 2002/2003 im geschichtlichen und waldökologischen Kontext*. Grafenau, 178 S
- Hikosaka, K. (2004). Interspecific difference in the photosynthesis–nitrogen relationship: patterns, physiological causes, and ecological importance. *Journal of Plant Research*, 117, 481-494
- Hogberg, P. (2007). Environmental science - Nitrogen impacts on forest carbon. *Nature*, 447, 781-782
- Homolova, L., Maenovsky, Z., Clevers, J., Garcia-Santos, G., & Schaepman, M.E. (2013). Review of optical-based remote sensing for plant trait mapping. *Ecological Complexity*, 15, 1-16
- Hosgood, B., Jacquemoud, S., Andreoli, G., Verdebout, J.P.G., & Schmuck, G. (1995). Leaf Optical Properties Experiment 93 (LOPEX93) Report EUR-16095-EN. *Ispira (Italy) European Commission—Joint Research Centre*
- Houborg, R., Anderson, M., & Daughtry, C. (2009). Utility of an image-based canopy reflectance modeling tool for remote estimation of LAI and leaf chlorophyll content at the field scale. *Remote Sensing of Environment*, 113, 259-274

- Houborg, R., McCabe, M., Cescatti, A., Gao, F., Schull, M., & Gitelson, A. (2015). Joint leaf chlorophyll content and leaf area index retrieval from Landsat data using a regularized model inversion system (REGFLEC). *Remote Sensing of Environment*, 159, 203-221
- Huang, D., Knyazikhin, Y., Dickinson, R.E., Rautiainen, M., Stenberg, P., Disney, M., Lewis, P., Cescatti, A., Tian, Y.H., Verhoef, W., Martonchik, J.V., & Myneni, R.B. (2007). Canopy spectral invariants for remote sensing and model applications. *Remote Sensing of Environment*, 106, 106-122
- Huang, Z., Turner, B.J., Dury, S.J., Wallis, I.R., & Foley, W.J. (2004). Estimating foliage nitrogen concentration from HYMAP data using continuum removal analysis. *Remote Sensing of Environment*, 93, 18-29
- Huemmrich, K.F. (2001). The GeoSail model: a simple addition to the SAIL model to describe discontinuous canopy reflectance. *Remote Sensing of Environment*, 75, 423-431
- Inoue, Y., Sakaiya, E., Zhu, Y., & Takahashi, W. (2012). Diagnostic mapping of canopy nitrogen content in rice based on hyperspectral measurements. *Remote Sensing of Environment*, 126, 210-221
- Jacquemoud, S. (1993). Inversion of the PROSPECT + SAIL canopy reflectance model from AVIRIS equivalent spectra: Theoretical study. *Remote Sensing of Environment*, 44, 281-292
- Jacquemoud, S., Bacour, C., Poilve, H., & Frangi, J.P. (2000). Comparison of four radiative transfer models to simulate plant canopies reflectance: Direct and inverse mode. *Remote Sensing of Environment*, 74, 471-481
- Jacquemoud, S., & Baret, F. (1990). PROSPECT: A model of leaf optical properties spectra. *Remote Sensing of Environment*, 34, 75-91
- Jacquemoud, S., Ustin, S.L., Verdebout, J., Schmuck, G., Andreoli, G., & Hosgood, B. (1996). Estimating leaf biochemistry using the PROSPECT leaf optical properties model. *Remote Sensing of Environment*, 56, 194-202
- Jacquemoud, S., Verdebout, J., Schmuck, G., Andreoli, G., & Hosgood, B. (1995). Investigation of leaf biochemistry by statistics. *Remote Sensing of Environment*, 54, 180-188
- Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P.J., Asner, G.P., Francois, C., & Ustin, S.L. (2009). PROSPECT plus SAIL models: A review of use for vegetation characterization. *Remote Sensing of Environment*, 113, S56-S66
- Jurdao, S., Yebra, M., Guerschman, J.P., & Chuvieco, E. (2013). Regional estimation of woodland moisture content by inverting Radiative Transfer Models. *Remote Sensing of Environment*, 132, 59-70

- Kaewpijit, S., Le Moigne, J., & El-Ghazawi, T. (2003). Automatic reduction of hyperspectral imagery using wavelet spectral analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 41, 863-871
- Kaufmann, H., Segl, K., Guanter, L., Hofer, S., Foerster, K.P., Stuffer, T., Mueller, A., Richter, R., Bach, H., Hostert, P., & Chlebek, C. (2008). Environmental mapping and analysis program (EnMAP) - recent advances and status. In, *IEEE International Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008*. (pp. IV - 109-IV - 112)
- Kimes, D.S., Knyazikhin, Y., Privette, J.L., Abuelgasim, A.A., & Gao, F. (2000). Inversion methods for physically-based models. *Remote Sensing Reviews*, 18, 381-439
- Knyazikhin, Y., Schull, M.A., Stenberg, P., Mottus, M., Rautiainen, M., Yang, Y., Marshak, A., Latorre Carmona, P., Kaufmann, R.K., Lewis, P., Disney, M.I., Vanderbilt, V., Davis, A.B., Baret, F., Jacquemoud, S., Lyapustin, A., & Myneni, R.B. (2013a). Hyperspectral remote sensing of foliar nitrogen content. *Proceedings of the National Academy of Sciences of the United States of America*, 110, E185-192
- Knyazikhin, Y., Lewis, P., Disney, M.I., Stenberg, P., Mottus, M., Rautiainen, M., Kaufmann, R.K., Marshak, A., Schull, M.A., Carmona, P.L., Vanderbilt, V., Davis, A.B., Baret, F., Jacquemoud, S., Lyapustin, A., Yang, Y., & Myneni, R.B. (2013b). Reply to Townsend et al.: Decoupling contributions from canopy structure and leaf optics is critical for remote sensing leaf biochemistry. *Proceedings of the National Academy of Sciences of the United States of America*, 110, E1075-E1075
- Knyazikhin, Y., Schull, M.A., Xu, L., Myneni, R.B., & Samanta, A. (2011). Canopy spectral invariants. Part 1: A new concept in remote sensing of vegetation. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 112, 727-735
- Kokaly, R.F. (2001). Investigating a physical basis for spectroscopic estimates of leaf nitrogen concentration. *Remote Sensing of Environment*, 75, 153-161
- Kokaly, R.F., Asner, G.P., Ollinger, S.V., Martin, M.E., & Wessman, C.A. (2009). Characterizing canopy biochemistry from imaging spectroscopy and its application to ecosystem studies. *Remote Sensing of Environment*, 113, S78-S91
- Kokaly, R.F., & Clark, R.N. (1999). Spectroscopic determination of leaf biochemistry using band-depth analysis of absorption features and stepwise multiple linear regression. *Remote Sensing of Environment*, 67, 267-287

- Kou, L., Labrie, D., & Chylek, P. (1993). Refractive indices of water and ice in the 0.65- to 2.5 micrometer spectral range. *Applied Optics*, 32, 3531-3540
- Kramer, P.J., & Boyer, J.S. (1995). *Water Relations of Plants and Soils*. Academic Press
- LaCapra, V.C., Melack, J.M., Gastil, M., & Valeriano, D. (1996). Remote sensing of foliar chemistry of inundated rice with imaging spectrometry. *Remote Sensing of Environment*, 55, 50-58
- Lamarque, J.F., Kiehl, J.T., Brasseur, G.P., Butler, T., Cameron-Smith, P., Collins, W.D., Collins, W.J., Granier, C., Hauglustaine, D., Hess, P.G., Holland, E.A., Horowitz, L., Lawrence, M.G., McKenna, D., Merilees, P., Prather, M.J., Rasch, P.J., Rotman, D., Shindell, D., & Thornton, P. (2005a). Assessing future nitrogen deposition and carbon cycle feedback using a multimodel approach: Analysis of nitrogen deposition. *Journal of Geophysical Research-Atmospheres*, 110
- Lamarque, J.F., Kiehl, J.T., Brasseur, G.P., Butler, T., Cameron-Smith, P., Collins, W.D., Collins, W.J., Granier, C., Hauglustaine, D., Hess, P.G., Holland, E.A., Horowitz, L., Lawrence, M.G., McKenna, D., Merilees, P., Prather, M.J., Rasch, P.J., Rotman, D., Shindell, D., & Thornton, P. (2005d). Assessing future nitrogen deposition and carbon cycle feedback using a multimodel approach: Analysis of nitrogen deposition. *Journal of Geophysical Research-Atmospheres*, 110, 1-21
- Landis, R.S. (2005). Standardized Regression Coefficients. *Encyclopedia of Statistics in Behavioral Science*: John Wiley & Sons, Ltd
- Latorre-Carmona, P., Knyazikhin, Y., Alonso, L., Moreno, J.F., Pla, F., & Yan, Y. (2014). On Hyperspectral Remote Sensing of Leaf Biophysical Constituents: Decoupling Vegetation Structure and Leaf Optics Using CHRIS-PROBA Data Over Crops in Barrax. *IEEE Geoscience and Remote Sensing Letters*, 11, 1579-1583
- Laurent, V.C.E., Verhoef, W., Clevers, J., & Schaepman, M.E. (2011). Inversion of a coupled canopy-atmosphere model using multi-angular top-of-atmosphere radiance data: A forest case study. *Remote Sensing of Environment*, 115, 2603-2612
- Lausch, A., Heurich, M., & Fahse, L. (2013). Spatio-temporal infestation patterns of *Ips typographus* (L.) in the Bavarian Forest National Park, Germany. *Ecological Indicators*, 31, 73-81
- le Maire, G., Francois, C., & Dufrene, E. (2004). Towards universal broad leaf chlorophyll indices using PROSPECT simulated database and

- hyperspectral reflectance measurements. *Remote Sensing of Environment*, 89, 1-28
- le Maire, G., Francois, C., Soudani, K., Berveiller, D., Pontauiller, J.Y., Breda, N., Genet, H., Davi, H., & Dufrene, E. (2008). Calibration and validation of hyperspectral indices for the estimation of broadleaved forest leaf chlorophyll content, leaf mass per area, leaf area index and leaf canopy biomass. *Remote Sensing of Environment*, 112, 3846-3864
- LeBauer, D.S., & Treseder, K.K. (2008). Nitrogen limitation of net primary productivity in terrestrial ecosystems is globally distributed. *Ecology*, 89, 371-379
- Leblanc, S.G., Chen, J.M., Fernandes, R., Deering, D.W., & Conley, A. (2005). Methodology comparison for canopy structure parameters extraction from digital hemispherical photography in boreal forests. *Agricultural and Forest Meteorology*, 129, 187-207
- Lepine, L.C., Ollinger, S.V., Ouimette, A.P., & Martin, M.E. (2016). Examining spectral reflectance features related to foliar nitrogen in forests: Implications for broad-scale nitrogen mapping. *Remote Sensing of Environment*, 173, 174-186
- Lewis, P., & Disney, M. (2007). Spectral invariants and scattering across multiple scales from within-leaf to canopy. *Remote Sensing of Environment*, 109, 196-206
- Li, F., Miao, Y., Feng, G., Yuan, F., Yue, S., Gao, X., Liu, Y., Liu, B., Ustin, S.L., & Chen, X. (2014). Improving estimation of summer maize nitrogen status with red edge-based spectral vegetation indices. *Field Crops Research*, 157, 111-123
- Li, P.H., & Wang, Q. (2011). Retrieval of leaf biochemical parameters using PROSPECT inversion: A new approach for alleviating ill-posed problems. *IEEE Transactions on Geoscience and Remote Sensing*, 49, 2499-2506
- Li, X., & Strahler, A. (1985). Geometric-optical modeling of a conifer forest canopy. *IEEE Transactions on Geoscience and Remote Sensing*, GE-23, 705-721
- Li, X., & Strahler, A.H. (1992). Geometric-optical bidirectional reflectance modeling of the discrete crown vegetation canopy - effect of crown shape and mutual shadowing. *IEEE Transactions on Geoscience and Remote Sensing*, 30, 276-292
- Liang, L., Qin, Z., Zhao, S., Di, L., Zhang, C., Deng, M., Lin, H., Zhang, L., Wang, L., & Liu, Z. (2016). Estimating crop chlorophyll content with

- hyperspectral vegetation indices and the hybrid inversion method. *International Journal of Remote Sensing*, 37, 2923-2949
- Liu, J.X., Price, D.T., & Chen, J.A. (2005). Nitrogen controls on ecosystem carbon sequestration: a model implementation and application to Saskatchewan, Canada. *Ecological Modelling*, 186, 178-195
- Lloyd, J., Bloomfield, K., Domingues, T.F., & Farquhar, G.D. (2013). Photosynthetically relevant foliar traits correlating better on a mass vs an area basis: of ecophysiological relevance or just a case of mathematical imperatives and statistical quicksand? *New Phytologist*, 199, 311-321
- Long, S.P. (1991). Modification of the response of photosynthetic productivity to rising temperature by atmospheric CO<sub>2</sub> concentrations: Has its importance been underestimated? *Plant, Cell & Environment*, 14, 729-739
- Ma, Z.G., Chen, X., Wang, Q., Li, P.H., & Jiapaerl, G. (2012). Retrieval of leaf biochemical properties by inversed PROSPECT model and hyperspectral indices: an application to *Populus euphratica* polymorphic leaves. *Journal of Arid Land*, 4, 52-62
- Macfarlane, C. (2011). Classification method of mixed pixels does not affect canopy metrics from digital images of forest overstorey. *Agricultural and Forest Meteorology*, 151, 833-840
- Magnani, F., Mencuccini, M., Borghetti, M., Berbigier, P., Berninger, F., Delzon, S., Grelle, A., Hari, P., Jarvis, P.G., Kolari, P., Kowalski, A.S., Lankreijer, H., Law, B.E., Lindroth, A., Loustau, D., Manca, G., Moncrieff, J.B., Rayment, M., Tedeschi, V., Valentini, R., & Grace, J. (2007). The human footprint in the carbon cycle of temperate and boreal forests. *Nature*, 447, 848-850
- Majeke, B., van Aardt, J.A.N., & Cho, M.A. (2008). Imaging spectroscopy of foliar biochemistry in forestry environments. *Southern Forests*, 70, 275-285
- Malenovsky, Z., Albrechtova, J., Lhotakova, Z., Zurita-Milla, R., Clevers, J.G.P.W., Schaepman, M.E., & Cudlin, P. (2006). Applicability of the PROSPECT model for Norway spruce needles. *International Journal of Remote Sensing*, 27, 5315-5340
- Malenovsky, Z., Homolova, L., Zurita-Milla, R., Lukes, P., Kaplan, V., Hanus, J., Gastellu-Etchegorry, J.-P., & Schaepman, M.E. (2013). Retrieval of spruce leaf chlorophyll content from airborne image data using continuum removal and radiative transfer. *Remote Sensing of Environment*, 131, 85-102



- Malenovsky, Z., Rott, H., Cihlar, J., Schaepman, M.E., Garcia-Santos, G., Fernandes, R., & Berger, M. (2012). Sentinels for science: Potential of Sentinel-1, -2, and -3 missions for scientific observations of ocean, cryosphere, and land. *Remote Sensing of Environment*, 120, 91-101
- Mallat, S.G. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11, 674-693
- Markwell, J., Osterman, J., & Mitchell, J. (1995). Calibration of the Minolta SPAD-502 leaf chlorophyll meter. *Photosynthesis Research*, 46, 467-472
- Marten, G., Shenk, J., & Barton, F., II (1985). Near infrared reflectance spectroscopy (NIRS): Analysis of forage quality. *Agriculture Handbook No. 643*, 643
- Martin, M.E., & Aber, J.D. (1997). High spectral resolution remote sensing of forest canopy lignin, nitrogen, and ecosystem processes. *Ecological Applications*, 7, 431-443
- Martin, M.E., Plourde, L.C., Ollinger, S.V., Smith, M.L., & McNeil, B.E. (2008). A generalizable method for remote sensing of canopy nitrogen across a wide range of forest ecosystems. *Remote Sensing of Environment*, 112, 3511-3519
- McNeil, B.E., Read, J.M., & Driscoll, C.T. (2007). Foliar nitrogen responses to elevated atmospheric nitrogen deposition in nine temperate forest canopy species. *Environmental Science & Technology*, 41, 5191-5197
- McNeil, B.E., Read, J.M., Sullivan, T.J., McDonnell, T.C., Fernandez, I.J., & Driscoll, C.T. (2008). The spatial pattern of nitrogen cycling in the Adirondack Park, New York. *Ecological Applications*, 18, 438-452
- Meher, S.K., Shankar, B.U., & Ghosh, A. (2007). Wavelet-Feature-Based Classifiers for Multispectral Remote-Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing*, 45, 1881-1886
- Meroni, M., Colombo, R., & Panigada, C. (2004). Inversion of a radiative transfer model with hyperspectral observations for LAI mapping in poplar plantations. *Remote Sensing of Environment*, 92, 195-206
- Meziane, D., & Shipley, B. (1999). Interacting determinants of specific leaf area in 22 herbaceous species: effects of irradiance and nutrient availability. *Plant, Cell & Environment*, 22, 447-459
- Miller, J.R., Hare, E.W., & Wu, J. (1990). Quantitative characterization of the vegetation red edge reflectance 1. An inverted-Gaussian reflectance model. *International Journal of Remote Sensing*, 11, 1755-1773

- Miphokasap, P., Honda, K., Vaiphasa, C., Souris, M., & Nagai, M. (2012). Estimating Canopy Nitrogen Concentration in Sugarcane Using Field Imaging Spectroscopy. *Remote Sensing*, 4, 1651-1670
- Moorthy, I., Miller, J.R., & Noland, T.L. (2008). Estimating chlorophyll concentration in conifer needles with hyperspectral data: An assessment at the needle and canopy level. *Remote Sensing of Environment*, 112, 2824-2838
- Morford, S.L., Houlton, B.Z., & Dahlgren, R.A. (2011). Increased forest ecosystem carbon and nitrogen storage from nitrogen rich bedrock. *Nature*, 477, 78-81
- Mutanga, O., & Skidmore, A.K. (2004). Integrating imaging spectroscopy and neural networks to map grass quality in the Kruger National Park, South Africa. *Remote Sensing of Environment*, 90, 104-115
- Mutanga, O., & Skidmore, A.K. (2007). Red edge shift and biochemical content in grass canopies. *ISPRS Journal of Photogrammetry and Remote Sensing*, 62, 34-42
- Mutanga, O., Skidmore, A.K., Kumar, L., & Ferwerda, J. (2005). Estimating tropical pasture quality at canopy level using band depth analysis with continuum removal in the visible domain. *International Journal of Remote Sensing*, 26, 1093-1108
- Mutanga, O., Skidmore, A.K., & Prins, H.H.T. (2004). Predicting in situ pasture quality in the Kruger National Park, South Africa, using continuum-removed absorption features. *Remote Sensing of Environment*, 89, 393-408
- Mutanga, O., Skidmore, A.K., & Van Wieren, S. (2003). Discriminating tropical grass (*Cenchrus ciliaris*) canopies grown under different nitrogen treatments using spectroradiometry. *ISPRS Journal of Photogrammetry and Remote Sensing*, 57, 263-272
- Neter, J., Wasserman, W., & Kutner, M.H. (1990). *Applied Linear Statistical Models*. (3rd ed.). Irwin, Boston, MA
- Niemann, K.O., Quinn, G., Goodenough, D.G., Visintini, F., & Loos, R. (2012). Addressing the effects of canopy structure on the remote sensing of foliar chemistry of a 3-dimensional, radiometrically porous surface. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5, 584-593
- Niinemets, U. (2010). A review of light interception in plant stands from leaf to canopy in different plant functional types and in species with varying shade tolerance. *Ecological Research*, 25, 693-714

- Niinemets, Ü. (1997). Role of foliar nitrogen in light harvesting and shade tolerance of four temperate deciduous woody species. *Functional Ecology*, *11*, 518-531
- Norris, K.H., Barnes, R.F., Moore, J.E., & Shenk, J.S. (1976). Predicting forage quality by infrared reflectance spectroscopy. *Journal of Animal Science*, *43*, 889-897
- Novozamsky, I., Houba, V.J.G., van Eck, R., & van Vark, W. (1983). A novel digestion technique for multi-element plant analysis. *Communications in Soil Science and Plant Analysis*, *14*, 239-248
- (NRC), N.R.C. (2007). Earth science and applications from space: national imperatives for the next decade and beyond. Washington, DC, USA: National Academies Press
- O'Brien, R.M. (2007). A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality & Quantity*, *41*, 673-690
- Ollinger, S. (2011). Sources of variability in canopy reflectance and the convergent properties of plants. *New Phytologist*, *189*, 375-394
- Ollinger, S.V., Reich, P.B., Frolking, S., Lepine, L.C., Hollinger, D.Y., & Richardson, A.D. (2013). Nitrogen cycling, forest canopy reflectance, and emergent properties of ecosystems. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, E2437-E2437
- Ollinger, S.V., Richardson, A.D., Martin, M.E., Hollinger, D.Y., Frolking, S.E., Reich, P.B., Plourde, L.C., Katul, G.G., Munger, J.W., Oren, R., Smithb, M.L., U, K.T.P., Bolstad, P.V., Cook, B.D., Day, M.C., Martin, T.A., Monson, R.K., & Schmid, H.P. (2008). Canopy nitrogen, carbon assimilation, and albedo in temperate and boreal forests: Functional relations and potential climate feedbacks. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 19336-19341
- Ollinger, S.V., & Smith, M.L. (2005). Net primary production and canopy nitrogen in a temperate forest landscape: An analysis using imaging spectroscopy, modeling and field data. *Ecosystems*, *8*, 760-778
- Omari, K., White, H.P., Staenz, K., & King, D.J. (2013). Retrieval of forest canopy parameters by inversion of the PROFLAIR leaf-canopy reflectance model using the LUT approach. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *6*, 715-723
- Oppelt, N., & Mauser, W. (2004). Hyperspectral monitoring of physiological parameters of wheat during a vegetation period using AVIS data. *International Journal of Remote Sensing*, *25*, 145-159

- Osborne, B.G., & Fearn, T. (1986). *Near Infrared Spectroscopy in Food Analysis*. London: Longman
- Otazu, X., Gonzalez-Audicana, M., Fors, O., & Nunez, J. (2005). Introduction of sensor spectral response into image fusion methods. Application to wavelet-based methods. *IEEE Transactions on Geoscience and Remote Sensing*, 43, 2376-2385
- Pellissier, P.A., Ollinger, S.V., Lepine, L.C., Palace, M.W., & McDowell, W.H. (2015). Remote sensing of foliar nitrogen in cultivated grasslands of human dominated landscapes. *Remote Sensing of Environment*, 167, 88-97
- Pereira, H.M., Ferrier, S., Walters, M., Geller, G.N., Jongman, R.H.G., Scholes, R.J., Bruford, M.W., Brummitt, N., Butchart, S.H.M., Cardoso, A.C., Coops, N.C., Dullo, E., Faith, D.P., Freyhof, J., Gregory, R.D., Heip, C., Höft, R., Hurtt, G., Jetz, W., Karp, D.S., McGeoch, M.A., Obura, D., Onoda, Y., Pettorelli, N., Reyers, B., Sayre, R., Scharlemann, J.P.W., Stuart, S.N., Turak, E., Walpole, M., & Wegmann, M. (2013). Essential biodiversity variables. *Science*, 339, 277-278
- Peterson, D.L., Aber, J.D., Matson, P.A., Card, D.H., Swanberg, N., Wessman, C., & Spanner, M. (1988). Remote sensing of forest canopy and leaf biochemical contents. *Remote Sensing of Environment*, 24, 85-108
- Peterson, D.L., & Hubbard, G.S. (1992). Scientific issues and potential remote-sensing requirements for plant biochemical content. *Journal of Imaging Science and Technology*, 36, 446-456
- Pinty, B., Widlowski, J.L., Taberner, M., Gobron, N., Verstraete, M.M., Disney, M., Gascon, F., Gastellu, J.P., Jiang, L., Kuusk, A., Lewis, P., Li, X., Ni-Meister, W., Nilson, T., North, P., Qin, W., Su, L., Tang, S., Thompson, R., Verhoef, W., Wang, H., Wang, J., Yan, G., & Zang, H. (2004). Radiation Transfer Model Intercomparison (RAMI) exercise: Results from the second phase. *Journal of Geophysical Research-Atmospheres*, 109, D06210
- Plummer, S.E. (2000). Perspectives on combining ecological process models and remotely sensed data. *Ecological Modelling*, 129, 169-186
- Poorter, L., Oberbauer, S.F., & Clark, D.B. (1995). Leaf optical properties along a vertical gradient in a tropical rain forest canopy in Costa Rica. *American Journal of Botany*, 1257-1263
- Porra, R. (2002). The chequered history of the development and use of simultaneous equations for the accurate determination of chlorophylls a and b. *Photosynthesis Research*, 73, 149-156

- Pragnere, A., Baret, F., Weiss, M., Myneni, R., Knyazikhin, Y., & Wang, L.B. (1999). Comparison of three radiative transfer model inversion techniques to estimate canopy biophysical variables from remote sensing data. In, *IEEE Geoscience and Remote Sensing Symposium, 1999. IGARSS '99 Proceedings*, 1093-1095
- Ramoelo, A., Skidmore, A.K., Schlerf, M., Mathieu, R., & Heitkönig, I.M.A. (2011). Water-removed spectra increase the retrieval accuracy when estimating savanna grass nitrogen and phosphorus concentrations. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66, 408-417
- Reich, P.B. (2012). Key canopy traits drive forest productivity. *Proceedings of the Royal Society B: Biological Sciences*, 279, 2128-2134
- Reich, P.B., Ellsworth, D.S., & Walters, M.B. (1998). Leaf structure (specific leaf area) modulates photosynthesis–nitrogen relations: evidence from within and across species and functional groups. *Functional Ecology*, 12, 948-958
- Reich, P.B., Hobbie, S.E., Lee, T., Ellsworth, D.S., West, J.B., Tilman, D., Knops, J.M.H., Naeem, S., & Trost, J. (2006). Nitrogen limitation constrains sustainability of ecosystem response to CO<sub>2</sub>. *Nature*, 440, 922-925
- Reich, P.B., Walters, M.B., & Ellsworth, D.S. (1997). From tropics to tundra: Global convergence in plant functioning. *Proceedings of the National Academy of Sciences of the United States of America*, 94, 13730-13734
- Riano, D., Vaughan, P., Chuvieco, E., Zarco-Tejada, P.J., & Ustin, S.L. (2005). Estimation of fuel moisture content by inversion of radiative transfer models to simulate equivalent water thickness and dry matter content: analysis at leaf and canopy level. *IEEE Transactions on Geoscience and Remote Sensing*, 43, 819-826
- Rogge, D.M., & Rivard, B. (2010). Iterative spatial filtering for reducing intra-class spectral variability and noise. In, *Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2010 2nd Workshop on* (pp. 1-4)
- Romero, A., Aguado, I., & Yebra, M. (2012). Estimation of dry matter content in leaves using normalized indexes and PROSPECT model inversion. *International Journal of Remote Sensing*, 33, 396-414
- Rosati, A., Day, K.R., & DeJong, T.M. (2000). Distribution of leaf mass per unit area and leaf nitrogen concentration determine partitioning of leaf nitrogen within tree canopies. *Tree Physiology*, 20, 271-276

- Rosema, A., Verhoef, W., Noorbergen, H., & Borgesius, J.J. (1992). A new forest light interaction model in support of forest monitoring. *Remote Sensing of Environment*, 42, 23-41
- Ross, I.U. (1981). *The radiation regime and architecture of plant stands*. Springer
- Saltelli, A. (1999). Sensitivity analysis: Could better methods be used? *Journal of Geophysical Research: Atmospheres*, 104, 3789-3793
- Saltelli, A., & Bolado, R. (1998). An alternative way to compute Fourier amplitude sensitivity test (FAST). *Computational Statistics & Data Analysis*, 26, 445-460
- Savitzky, A., & Golay, M.J.E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 36, 1627-1639
- Schläpfer, D., & Richter, R. (2011). Spectral Polishing of High Resolution Imaging Spectroscopy Data. In, *7th SIG-IS Workshop on Imaging Spectroscopy* (pp. 1-7). Edinburgh
- Schlemmer, M., Gitelson, A.A., Schepers, J., Ferguson, R., Peng, Y., Shanahan, J., & Rundquist, D. (2013). Remote estimation of nitrogen and chlorophyll contents in maize at leaf and canopy levels. *International Journal of Applied Earth Observation and Geoinformation*, 25, 47-54
- Schlerf, M., & Atzberger, C. (2006). Inversion of a forest reflectance model to estimate structural canopy variables from hyperspectral remote sensing data. *Remote Sensing of Environment*, 100, 281-294
- Schlerf, M., & Atzberger, C. (2012). Vegetation structure retrieval in beech and spruce forests using spectrodirectional satellite data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5, 8-17
- Schlerf, M., Atzberger, C., & Hill, J. (2005). Remote sensing of forest biophysical variables using HyMap imaging spectrometer data. *Remote Sensing of Environment*, 95, 177-194
- Schlerf, M., Atzberger, C., Hill, J., Buddenbaum, H., Werner, W., & Schueler, G. (2010). Retrieval of chlorophyll and nitrogen in Norway spruce (*Picea abies* L. Karst.) using imaging spectroscopy. *International Journal of Applied Earth Observation and Geoinformation*, 12, 17-26
- Schmidt, K.S., & Skidmore, A.K. (2004). Smoothing vegetation spectra with wavelets. *International Journal of Remote Sensing*, 25, 1167-1184
- Schull, M.A., Knyazikhin, Y., Xu, L., Samanta, A., Carmona, P.L., Lepine, L., Jenkins, J.P., Ganguly, S., & Myneni, R.B. (2011). Canopy spectral invariants, Part 2: Application to classification of forest types from

- hyperspectral data. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 112, 736-750
- Scott Green, D., Erickson, J.E., & Kruger, E.L. (2003). Foliar morphology and canopy nitrogen as predictors of light-use efficiency in terrestrial vegetation. *Agricultural and Forest Meteorology*, 115, 163-171
- Serrano, L., Peñuelas, J., & Ustin, S.L. (2002). Remote sensing of nitrogen and lignin in Mediterranean vegetation from AVIRIS data: Decomposing biochemical from structural signals. *Remote Sensing of Environment*, 81, 355-364
- Shi, W., Zhu, C., Tian, Y., & Nichol, J. (2005). Wavelet-based image fusion and quality assessment. *International Journal of Applied Earth Observation and Geoinformation*, 6, 241-251
- Si, Y., Schlerf, M., Zurita-Milla, R., Skidmore, A., & Wang, T. (2012). Mapping spatio-temporal variation of grassland quantity and quality using MERIS data and the PROSAIL model. *Remote Sensing of Environment*, 121, 415-425
- Sievering, H., Fernandez, I., Lee, J., Hom, J., & Rustad, L. (2000). Forest canopy uptake of atmospheric nitrogen deposition at eastern US conifer sites: Carbon storage implications? *Global Biogeochemical Cycles*, 14, 1153-1159
- Singh, A., Serbin, S.P., McNeil, B.E., Kingdon, C.C., & Townsend, P.A. (2015). Imaging spectroscopy algorithms for mapping canopy foliar chemical and morphological traits and their uncertainties. *Ecological Applications*
- Skidmore, A.K. (2002). Taxonomy of environmental models in the spatial sciences. In A.K. Skidmore (Ed.), *Environmental modelling with GIS and remote sensing* (pp. 2-7). London: Taylor & Francis
- Skidmore, A.K., Ferwerda, J.G., Mutanga, O., Van Wieren, S.E., Peel, M., Grant, R.C., Prins, H.H.T., Balcik, F.B., & Venus, V. (2010). Forage quality of savannas — Simultaneously mapping foliar protein and polyphenols for trees and grass using hyperspectral imagery. *Remote Sensing of Environment*, 114, 64-72
- Skidmore, A.K., Pettorelli, N., Coops, N.C., Geller, G.N., Hansen, M., Lucas, R., Múcher, C.A., O'Connor, B., Paganini, M., Pereira, H.M., Schaepman, M.E., Turner, W., Wang, T., & Wegmann, M. (2015). Environmental science: Agree on biodiversity metrics to track from space. *Nature*, 523, 403-405
- Smith, M.L., Martin, M.E., Plourde, L., & Ollinger, S.V. (2003). Analysis of hyperspectral data for estimation of temperate forest canopy nitrogen

- concentration: comparison between an airborne (AVIRIS) and a spaceborne (Hyperion) sensor. *IEEE Transactions on Geoscience and Remote Sensing*, 41, 1332-1337
- Smith, M.L., Ollinger, S.V., Martin, M.E., Aber, J.D., Hallett, R.A., & Goodale, C.L. (2002). Direct estimation of aboveground forest productivity through hyperspectral remote sensing of canopy nitrogen. *Ecological Applications*, 12, 1286-1302
- Smolander, S., & Stenberg, P. (2005). Simple parameterizations of the radiation budget of uniform broadleaved and coniferous canopies. *Remote Sensing of Environment*, 94, 355-363
- Sobol, I.M. (1993). Sensitivity analysis for nonlinear mathematical models. *Mathematical Modeling and Computational Experiment*, 1, 407-414
- Steudler, P.A., Bowden, R.D., Melillo, J.M., & Aber, J.D. (1989). Influence of nitrogen fertilization on methane uptake in temperate forest soils. *Nature*, 341, 314-316
- Sullivan, F.B., Ollinger, S.V., Martin, M.E., Ducey, M.J., Lepine, L.C., & Wicklein, H.F. (2012). Foliar nitrogen in relation to plant traits and reflectance properties of New Hampshire forests. *Canadian Journal of Forest Research*, 43, 18-27
- Takahashi, K., Seino, T., & Kohyama, T. (2005). Plastic changes of leaf mass per area and leaf nitrogen content in response to canopy openings in saplings of eight deciduous broad-leaved tree species. *Ecological Research*, 20, 17-23
- Torrence, C., & Compo, G.P. (1998). A Practical Guide to Wavelet Analysis. *Bulletin of the American Meteorological Society*, 79, 61-78
- Townsend, P.A., Foster, J.R., Chastain, R.A., Jr., & Currie, W.S. (2003). Application of imaging spectroscopy to mapping canopy nitrogen in the forests of the central Appalachian Mountains using Hyperion and AVIRIS. *IEEE Transactions on Geoscience and Remote Sensing*, 41, 1347-1354
- Townsend, P.A., Serbin, S.P., Kruger, E.L., & Gamon, J.A. (2013). Disentangling the contribution of biological and physical properties of leaves and canopies in imaging spectroscopy data. *Proceedings of the National Academy of Sciences of the United States of America*, 110, E1074-E1074
- Ullah, S., Skidmore, A.K., Naeem, M., & Schlerf, M. (2012). An accurate retrieval of leaf water content from mid to thermal infrared spectra using continuous wavelet analysis. *Science of the Total Environment*, 437, 145-152



- Ustin, S.L. (2013). Remote sensing of canopy chemistry. *Proceedings of the National Academy of Sciences*, 110, 804-805
- Ustin, S.L., Jacquemoud, S., Zarco-Tejada, P.J., & Asner, G.P. (2004). Remote sensing of the environment: State of the science and new directions. *Manual of Remote Sensing*, 4, 679-730
- Ustin, S.L., Riano, D., & Hunt, E.R. (2012). Estimating canopy water content from spectroscopy. *Israel Journal of Plant Sciences*, 60, 9-23
- Van Soest, P.J., & Masson, V.C. (1967). The influence of the Maillard reaction upon the nutritive values of fibrous feed. *Feed Sci. Technol.*, 32, 45-53
- Van Soest, P.J., & Wine, R.H. (1991). Use of detergents in the analysis of fibrous feeds. IV. Determination of plant cell-wall constituents. *Journal of the Association of Official Analytical Chemists*, 50, 50-55
- Verhoef, W. (1984). Light scattering by leaf layers with application to canopy reflectance modeling: The SAIL model. *Remote Sensing of Environment*, 16, 125-141
- Verhoef, W., & Bach, H. (2007). Coupled soil-leaf-canopy and atmosphere radiative transfer modeling to simulate hyperspectral multi-angular surface reflectance and TOA radiance data. *Remote Sensing of Environment*, 109, 166-182
- Vitousek, P.M., Aber, J.D., Howarth, R.W., Likens, G.E., Matson, P.A., Schindler, D.W., Schlesinger, W.H., & Tilman, D.G. (1997). Human alteration of the global nitrogen cycle: sources and consequences. *Ecological Applications*, 7, 737-750
- Wang, Z., Skidmore, A.K., Wang, T., Darvishzadeh, R., & Hearne, J. (2015a). Applicability of the PROSPECT model for estimating protein and cellulose + lignin in fresh leaves. *Remote Sensing of Environment*, 168, 205-218
- Wang, Z., Skidmore, A.K., Darvishzadeh, R., Heiden, U., Heurich, M., & Wang, T. (2015b). Leaf nitrogen content indirectly estimated by leaf traits derived from the PROSPECT model. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8, 3172-3182
- Wang, Z., Wang, T., Darvishzadeh, R., Skidmore, A., Jones, S., Suarez, L., Woodgate, W., Heiden, U., Heurich, M., & Hearne, J. (2016a). Vegetation indices for mapping canopy foliar nitrogen in a mixed temperate forest. *Remote Sensing*, 8, 491
- Wang, Z., Skidmore, A.K., Wang, T., Darvishzadeh, R., Heiden, U., Heurich, M., Latifi, H., and Hearne J. (2016b). Canopy foliar nitrogen retrieved from airborne hyperspectral imagery by correcting for canopy structure effects.

- International Journal of Applied Earth Observation and Geoinformation*, 54, 84-94.
- Waring, R.H., & Landsberg, J.J. (2011). Generalizing plant–water relations to landscapes. *Journal of Plant Ecology*, 4, 101-113
- Weende, P. (1985). NF V03-400. *Recueil des Normes Françaises-AFNOR*
- Weiss, M., Baret, F., Myneni, R.B., Pragnère, A., & Knyazikhin, Y. (2000). Investigation of a model inversion technique to estimate canopy biophysical variables from spectral and directional reflectance data. *Agronomie*, 20, 3-22
- Wessman, C.A., Aber, J.D., Peterson, D.L., & Melillo, J.M. (1988). Remote sensing of canopy chemistry and nitrogen cycling in temperate forest ecosystems. *Nature*, 335, 154-156
- Widlowski, J.-L., Verstraete, M., Pinty, B., & Gobron, N. (2003). *Allometric relationships of selected european tree species: Parametrizations of tree architecture for the purpose of 3-D canopy reflectance models used in the interpretation of remote sensing: Betula pubescens, Fagus sylvatica, Larix decidua, Picea abies, Pinus sylvestris*. Ispra, Italy: EC Joint Research Centre
- Wieliczka, D.M., Weng, S., & Querry, M.R. (1989). Wedge shaped cell for highly absorbent liquids: Infrared optical constants of water. *Applied Optics*, 28, 1714-1719
- Witkowski, E.T.F., & Lamont, B. (1991). Leaf specific mass confounds leaf density and thickness. *Oecologia*, 88, 486-493
- Woodgate, W., Jones, S.D., Suarez, L., Hill, M.J., Armston, J.D., Wilkes, P., Soto-Berelev, M., Haywood, A., & Mellor, A. (2015). Understanding the variability in ground-based methods for retrieving canopy openness, gap fraction, and leaf area index in diverse forest systems. *Agricultural and Forest Meteorology*, 205, 83-95
- Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornellssen, J.H.C., Diemer, M., Flexas, J., Garnier, E., Groom, P.K., Gulias, J., Hikosaka, K., Lamont, B.B., Lee, T., Lee, W., Lusk, C., Midgley, J.J., Navas, M.L., Niinemets, Ü., Oleksyn, J., Osada, H., Poorter, H., Pool, P., Prior, L., Pyankov, V.I., Roumet, C., Thomas, S.C., Tjoelker, M.G., Veneklaas, E.J., & Villar, R. (2004). The worldwide leaf economics spectrum. *Nature*, 428, 821-827
- Xiao, Y., Zhao, W., Zhou, D., & Gong, H. (2014). Sensitivity analysis of vegetation reflectance to biochemical and biophysical variables at leaf,

- canopy, and regional scales. *IEEE Transactions on Geoscience and Remote Sensing*, 52, 4014-4024
- Yang, G.J., Zhao, C.J., Liu, Q., Huang, W.J., & Wang, J.H. (2011). Inversion of a radiative transfer model for estimating forest LAI From multisource and multiangular optical remote sensing data. *IEEE Transactions on Geoscience and Remote Sensing*, 49, 988-1000
- Yao, X., Huang, Y., Shang, G., Zhou, C., Cheng, T., Tian, Y., Cao, W., & Zhu, Y. (2015). Evaluation of six algorithms to monitor wheat leaf nitrogen concentration. *Remote Sensing*, 7, 14939-14966
- Yebra, M., & Chuvieco, E. (2009). Linking ecological information and radiative transfer models to estimate fuel moisture content in the Mediterranean region of Spain: Solving the ill-posed inverse problem. *Remote Sensing of Environment*, 113, 2403-2411
- Yebra, M., Chuvieco, E., & Riaño, D. (2008). Estimation of live fuel moisture content from MODIS images for fire risk assessment. *Agricultural and Forest Meteorology*, 148, 523-536
- Yoder, B.J., & Pettigrew-Crosby, R.E. (1995). Predicting nitrogen and chlorophyll content and concentrations from reflectance spectra (400–2500 nm) at leaf and canopy scales. *Remote Sensing of Environment*, 53, 199-211
- Yu, K. (2014). Hyperspectral Remote Sensing of Crop Canopy Chlorophyll and Nitrogen: The Relative Importance of Growth Stages. In (pp. 54-68): Universität zu Köln
- Yu, K., Li, F., Gnyp, M.L., Miao, Y.X., Bareth, G., & Chen, X.P. (2013). Remotely detecting canopy nitrogen concentration and uptake of paddy rice in the Northeast China Plain. *ISPRS Journal of Photogrammetry and Remote Sensing*, 78, 102-115
- Yuan, H., Ma, R., Atzberger, C., Li, F., Loisel, S., & Luo, J. (2015). Estimating forest fAPAR from multispectral Landsat-8 data using the invertible forest reflectance model INFORM. *Remote Sensing*, 7, 7425
- Zaehle, S., Medlyn, B.E., De Kauwe, M.G., Walker, A.P., Dietze, M.C., Hickler, T., Luo, Y., Wang, Y.-P., El-Masri, B., Thornton, P., Jain, A., Wang, S., Warland, D., Weng, E., Parton, W., Iversen, C.M., Gallet-Budynek, A., McCarthy, H., Finzi, A., Hanson, P.J., Prentice, I.C., Oren, R., & Norby, R.J. (2014). Evaluation of 11 terrestrial carbon–nitrogen cycle models against observations from two temperate Free-Air CO<sub>2</sub> Enrichment studies. *New Phytologist*, 202, 803-822

- Zarco-Tejada, P.J., Miller, J.R., Harron, J., Hu, B.X., Noland, T.L., Goel, N., Mohammed, G.H., & Sampson, P. (2004). Needle chlorophyll content estimation through model inversion using hyperspectral data from boreal conifer forest canopies. *Remote Sensing of Environment*, 89, 189-199
- Zarco-Tejada, P.J., Miller, J.R., Noland, T.L., Mohammed, G.H., & Sampson, P.H. (2001). Scaling-up and model inversion methods with narrowband optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 39, 1491-1507
- Zhang, C., Li, C., Chen, X., Luo, G., Li, L., Li, X., Yan, Y., & Shao, H. (2013). A spatial-explicit dynamic vegetation model that couples carbon, water, and nitrogen processes for arid and semiarid ecosystems. *Journal of Arid Land*, 5, 102-117
- Zhang, Y.Q., Chen, J.M., Miller, J.R., & Noland, T.L. (2008). Leaf chlorophyll content retrieval from airborne hyperspectral remote sensing imagery. *Remote Sensing of Environment*, 112, 3234-3247
- Zhao, F., Guo, Y., Huang, Y., Reddy, K.N., Lee, M.A., Fletcher, R.S., & Thomson, S.J. (2014). Early detection of crop injury from herbicide glyphosate by leaf biochemical parameter inversion. *International Journal of Applied Earth Observation and Geoinformation*, 31, 78-85
- Zhao, K., Valle, D., Popescu, S., Zhang, X., & Mallick, B. (2013). Hyperspectral remote sensing of plant biochemistry using Bayesian model averaging with variable and band selection. *Remote Sensing of Environment*, 132, 102-119
- Zhu, C., & Yang, X. (1998). Study of remote sensing image texture analysis and classification using wavelet. *International Journal of Remote Sensing*, 19, 3197-3203

## **Summary**

Foliar nitrogen is a critical parameter in leaf photosynthesis, respiration and transpiration, which relates to canopy and stand-level traits such as light use efficiency, wood growth and net primary production. Nitrogen plays an important role in terrestrial ecosystem carbon dynamics, and is an important input parameter of ecosystem process models. Recently, foliar nitrogen content has been proposed as one of the key essential biodiversity variables (EBVs) for satellite monitoring of progress towards the Aichi Biodiversity Targets.

Hyperspectral remote sensing provides a time-efficient and cost-effective solution to estimate foliar nitrogen. Empirical approaches have been the dominant when estimating nitrogen, and good accuracies achieved, though any empirical relationship so established is hard to be transferred across species, sites and sensors. Physically-based approach, or radiative transfer models (RTMs) simulate the process of radiation transfer based on physical laws, which is more robust and transferable. However, nitrogen has been considered to be impossible to be retrieved from fresh spectra by inversion of a leaf RTM model. Furthermore, at the canopy level, hyperspectral remote sensing of foliar nitrogen may be confounded by canopy structure which causes more variations in canopy bidirectional reflectance factor (BRF).

The thesis aimed to retrieve foliar nitrogen at the leaf and canopy level using radiative transfer models and with hyperspectral remote sensing. The research in this thesis firstly evaluated the feasibility of retrieving nitrogen by inversion of a leaf radiative transfer model PROSPECT. The indirect estimation of nitrogen by inversion of leaf PROSPECT through links between nitrogen and leaf traits was investigated, and it was proved that nitrogen can be estimated at a moderate accuracy by inversion of the leaf PROSPECT model and the empirical relationship between nitrogen and leaf traits (e.g. leaf water, leaf mass per area and chlorophyll). Secondly, the thesis recalibrated the specific absorption coefficients of protein and cellulose + lignin, and proved that the recalibrated model can accurately simulate leaf reflectance and transmittance. Nitrogen can be estimated at a moderate accuracy by inversion of the recalibrated model.

Then research moved on to the canopy level with emphasis on addressing canopy structural effects on nitrogen estimation. The thesis evaluated the feasibility of nitrogen retrieval from the canopy BRF in the full wavelengths between 425 -

2500 nm corrected for canopy structure effects. The results showed that nitrogen can still be detected from canopy BRF after correcting for canopy structure, and continuum wavelet analysis was shown to be an efficient method of revealing the nitrogen absorption features. The selected wavelet feature most sensitive to nitrogen variation were found to be correlated with water, dry matter and protein. Lastly, the thesis tested the capability of coupled leaf-canopy models in foliar nitrogen estimation. Global sensitivity analysis showed that the influence of nitrogen content on canopy reflectance is small but noticeable. Canopy nitrogen content was more accurately predicted than leaf nitrogen content. Maps of leaf and canopy nitrogen were generated for the study area, and spatial variations of nitrogen appeared consistent with ecological knowledge.

This thesis demonstrates an approach of estimating foliar nitrogen at the leaf and canopy level using leaf and canopy radiative transfer models. The robustness and transferability of radiative transfer models suggests that the approaches might be transferred to other sites. The combination of spectral invariants theory and continuous wavelet analysis have been shown to be an efficient and robust way of hyperspectral remote sensing of canopy foliar nitrogen regardless of the confounding effects of canopy structure. Accurate estimation of foliar nitrogen is beneficial in a variety of ecological applications. Global maps of foliar nitrogen content could be generated when hyperspectral satellites such as EnMAP and HypsIRI become operational. In turn, such monitoring would support the development and demonstration of Essential Biodiversity Variables (EBVs), which are used for assessing biodiversity. Regional and global mapping of foliar nitrogen will further our understanding of the photosynthesis process, net primary productivity, and carbon dynamics. In a broader context, foliar nitrogen as an input to ecosystem process models is essential in understanding the nutrient cycle and climatic feedbacks.

## Samenvatting

Blad stikstof is een belangrijke parameter als het gaat om de fotosynthese, ademhaling en transpiratie in bladeren. Deze hebben weer betrekking op de luifel niveau eigenschappen zoals licht-efficiëntie, de houtgroei en de netto primaire productie. Stikstof speelt een belangrijke rol in de koolstof dynamiek van het ruimtelijke ecosysteem. Ze zijn dan ook een belangrijke parameter voor de procesmodellen die gebruikt worden voor het ecosysteem. Blad stikstof wordt hedendaags ook gezien als een belangrijke variabele van de biodiversiteit (EBV) voor de voortgang van het monitoren van satellieten in de richting van de Aichi biodiversiteits doelen (Aichi Biodiversity Targets).

Het gebruik van hyperspectrale remote sensing geeft een efficiënte en kosten-effectieve oplossing om blad stikstof te voorspellen. Empirische benaderingen zijn dominant als het aankomt op het voorspellen van stikstof en een goede nauwkeurigheid is hierbij bereikt, alhoewel dit type empirische relaties moeilijk kunnen worden overgedragen tussen verschillende soorten, locaties en sensoren. Een fysieke benadering met modellen met stralingsoverdracht (RTM's) kunnen het stralingsproces simuleren gebaseerd op natuurkundige wetten, wat meer robuust en overdraagbaar is. Tot nu toe werd echter verwacht, dat het onmogelijk is om stikstof te verkrijgen vanuit nieuwe spectra door inversie van een blad-stralingsoverdracht (RTM). Daarnaast zou het mogelijk zijn dat, bij luifelhoogte, hyperspectrale remote sensing voor blad stikstof de luifel structuur zal verwarren, waardoor er meer variaties zullen ontstaan in de reflecterende factor voor het luifel (BRF).

Dit proefschrift tracht om blad stikstof te verkrijgen vanuit het blad alsmede de luifel hoogte door gebruik van modellen voor stralingsoverdracht (RTM) en hyperspectrale remote sensing. Het onderzoek in dit proefschrift zal eerst de mogelijkheid om stikstof te verkrijgen door de inversie van het PROSPECT stralingsmodel evalueren. Een indirecte schatting is gemaakt door het verband tussen stikstof en de blad eigenschappen te onderzoeken, en het is bewezen dat stikstof geschat kan worden met een bescheiden nauwkeurigheid door gebruik van het PROSPECT model tezamen met het empirische verband tussen stikstof en de bladeigenschappen (bijv. blad vocht, massa en chlorophyl aanwezigheid). Vervolgens zijn ook de specifieke absorptie coëfficiënten van proteïnen, cellulose en lignine geïk, en is zo bewezen dat het geïk model de blad reflectie en doorlaatbaarheid kan simuleren. Stikstof kan geschat worden, met een bescheiden nauwkeurigheid, door een inversie van het geïk model.

Vervolgens is het onderzoek gericht op de luifelhoogte met nadruk op het schatten van de structurele effecten op stikstof. De haalbaarheid van het verkrijgen van stikstof uit de luifel vanuit de golflengten tussen 425 en 2500 nm is onderzocht. De resultaten van het onderzoek laten zien dat stikstof nog steeds gedetecteerd kan worden na correctie van de luifel en in de analyse van de golflengtes bleek een goede methode om de absorptie van stikstof te onderzoeken. De gekozen golflengtes die het meest gevoelig waren voor stikstof bleken een correlatie te hebben met water, droog materiaal en eiwitten. Een algemene gevoeligheids analyse laat zien dat de stikstof op de luifelreflectie klein is maar wel aantoonbaar. De stikstof in de luifel kon nauwkeuriger voorspeld worden dan de stikstof in het blad. Er zijn kaarten gemaakt in het onderzoeksgebied van de aanwezigheid van stikstof in bladeren en luifel en de ruimtelijke variaties van de stikstof bleken consistent te zijn met de ecologische bevindingen.

Dit proefschrift laat de aanpak van het inschatten van stikstof in bladeren en op de luifel zien door gebruik van stralingsoverdracht modellen. De robuustheid en de overdraagzaamheid van de stralingsoverdrachts modellen suggereert dat deze aanpak ook gebruikt kan worden voor andere gebieden. Een nauwkeurige schatting van bladstikstof kan van belang zijn voor verschillende ecologische applicaties.

Wanneer hyperspectrale satellieten zoals EnMAP en HypsIRI operationeel worden, zouden er globale kaarten voor bladstikstof gegenereerd kunnen worden. Dit zou op zijn beurt weer de ontwikkeling van EBV's kunnen ondersteunen om op de biodiversiteit toe te passen. Regionale en Globale kaartering van bladstikstof kan ons begrip voor het proces van fotosynthese, de netto primaire productie en koolstof dynamiek bevorderen. In een breder context kan bladstikstof gebruikt worden als input in ecosysteem modellen wat essentieel is om de voedingscyclus en het klimaat beter te kunnen begrijpen.



## Biography



Zhihui Wang was born in Shandong, China on March 8, 1988. She received the BSc degree in geographical information science from Nanjing Normal University in 2009, and the ME degree in signal and information processing from Institute of Remote Sensing Applications, Chinese Academy of Sciences in 2012. In September, 2012, she was awarded a four-year doctoral scholarship from China Scholarship Council (CSC), and then started to pursue her PhD at the Faculty of Geo-Information

Science and Earth Observation (ITC), University of Twente. In July, 2013, she enrolled in a double-badged PhD degree program between University of Twente and RMIT University. Her research is focused on retrieval of leaf biochemical parameters using physically based models and hyperspectral remote sensing.

## **ITC Dissertation List**

[https://www.itc.nl/Pub/research\\_programme/Research-review-and-output/PhD-Graduates](https://www.itc.nl/Pub/research_programme/Research-review-and-output/PhD-Graduates)